

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2006, 20:25:27 ; Search time 180 Seconds
(without alignments)
1883.740 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826

Sequence: 1 LOEVHVSKEITKIGISASMKM.....CKCRDQWEGPHCENFLRRP 732

Scoring table: BLOSUM62

Gap: 10.0, Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA Main:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3826	100.0	732	4 US-10-229-345-3	Sequence 3, Appl1
2	3826	100.0	732	4 US-10-274-177-3	Sequence 3, Appl1
3	3826	100.0	732	4 US-10-650-112-3	Sequence 3, Appl1
4	3826	100.0	732	5 US-10-649-591-3	Sequence 3, Appl1
5	3826	100.0	732	6 US-11-041-788-3	Sequence 3, Appl1
6	3826	100.0	755	4 US-10-295-027-124	Sequence 124, App
7	3826	100.0	755	4 US-10-173-999-87	Sequence 87, Appl1
8	3826	100.0	755	4 US-10-229-345-14	Sequence 14, Appl1
9	3826	100.0	755	4 US-10-274-177-14	Sequence 14, Appl1
10	3826	100.0	755	4 US-10-650-112-14	Sequence 14, Appl1
11	3826	100.0	755	5 US-10-649-591-14	Sequence 14, Appl1
12	3826	100.0	755	6 US-11-041-788-14	Sequence 14, Appl1
13	3826	99.9	755	4 US-10-295-027-1186	Sequence 1186, Ap
14	3799	99.3	807	3 US-09-930-020A-2	Sequence 2, Appl1
15	3799	99.3	807	4 US-10-295-027-122	Sequence 122, App
16	3799	99.3	807	4 US-10-285-027-764	Sequence 764, App
17	3799	99.3	807	4 US-10-285-027-777	Sequence 777, App
18	3799	99.3	807	4 US-10-295-027-883	Sequence 883, App
19	3799	99.3	807	4 US-10-295-027-1187	Sequence 1187, Ap
20	3799	99.3	807	4 US-10-087-080-23	Sequence 23, Appl1
21	3799	99.3	807	4 US-10-173-999-85	Sequence 85, Appl1
22	3799	99.3	807	5 US-10-702-180-9	Sequence 2, Appl1
23	3387.5	88.5	1160	5 US-10-450-763-47702	Sequence 47702, A
24	3382.5	88.4	826	4 US-10-295-027-120	Sequence 120, App
25	3382.5	88.4	826	4 US-10-173-999-83	Sequence 83, Appl1
26	2568	67.1	488	5 US-10-649-591-21	Sequence 21, Appl1
27	2568	67.1	488	6 US-11-041-788-21	Sequence 21, Appl1

28	2562	67.0	488	4 US-10-650-112-21	Sequence 21, Appl1
29	1580	41.3	405	3 US-09-866-050A-374	Sequence 374, App
30	522.5	13.7	2211	5 US-10-741-600-1324	Sequence 1324, Ap
31	522.5	13.7	2244	5 US-10-741-600-1321	Sequence 1321, Ap
32	522.5	13.7	2411	5 US-10-741-600-1323	Sequence 1323, Ap
33	522.5	13.7	2977	5 US-10-741-600-1320	Sequence 1320, Ap
34	522.5	13.7	3151	5 US-10-486-678-16	Sequence 16, Appl1
35	522.5	13.7	3173	3 US-09-918-715-218	Sequence 218, App
36	522.5	13.7	3173	4 US-10-474-794-218	Sequence 218, App
37	522.5	13.7	3173	5 US-10-979-159-218	Sequence 218, App
38	522.5	13.7	3176	4 US-10-372-683-20	Sequence 20, Appl1
39	522.5	13.7	3176	4 US-10-734-564-91	Sequence 91, Appl1
40	522.5	13.7	3176	5 US-10-723-860-1065	Sequence 1065, Ap
41	522.5	13.7	3176	5 US-10-486-678-15	Sequence 15, Appl1
42	522.5	13.7	3176	5 US-10-852-335A-160	Sequence 160, App
43	522.5	13.7	3176	5 US-10-287-436A-503	Sequence 503, App
44	522.5	13.7	3176	5 US-10-287-436A-1196	Sequence 1196, App
45	522.5	13.7	3176	5 US-10-741-600-1319	Sequence 1319, Ap

ALIGNMENTS

RESULT 1
US-10-229-345-3
; Sequence 3, Application US//10229345
; Publication No. US20040038220A1
; GENERAL INFORMATION:
; APPLICANT: MARKOWITZ, Sanford D.
; TITLE OF INVENTION: METHODS FOR CATEGORIZING PATIENTS
; FILE REFERENCE: CMRU-P01-003
; CURRENT APPLICATION NUMBER: US/10/229,345
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-229-345-3

Query Match 100.0%; Score 3826; DB 4; Length 732;

Best Local Similarity 100.0%; Prod. No. 0;

Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LOEVHVSKEITKIGISASMKMCSAAVDIMFLDGSNSVKGKSFERSKHFAITVCDGLDI	60
DB	1	LOEVHVSKEITKIGISASMKMCSAAVDIMFLDGSNSVKGKSFERSKHFAITVCDGLDI	60
QY	61	SPERVVGAFOFSSTPHLEFPPLDSFSTQOEVARIKEMVFKGRTETELAKYLHARGUP	120
DB	61	SPERVVGAFOFSSTPHLEFPPLDSFSTQOEVARIKEMVFKGRTETELAKYLHARGUP	120
QY	121	GGNMSVPOILLIVTDGKSQGDVALPSKQKERGVTVFAVGVPRPWEELHAIASPRQO	180
DB	121	GGNMSVPOILLIVTDGKSQGDVALPSKQKERGVTVFAVGVPRPWEELHAIASPRQO	180
QY	121	GGNMSVPOILLIVTDGKSQGDVALPSKQKERGVTVFAVGVPRPWEELHAIASPRQO	180
DB	121	GGNMSVPOILLIVTDGKSQGDVALPSKQKERGVTVFAVGVPRPWEELHAIASPRQO	180
QY	181	HVLLAQVDATNGLFSTLSSAICGSATPDGCRVNAHPEHNTLBNVRFPAGNAPCWRS	240
DB	181	HVLLAQVDATNGLFSTLSSAICGSATPDGCRVNAHPEHNTLBNVRFPAGNAPCWRS	240
QY	181	HVLLAQVDATNGLFSTLSSAICGSATPDGCRVNAHPEHNTLBNVRFPAGNAPCWRS	240
DB	181	HVLLAQVDATNGLFSTLSSAICGSATPDGCRVNAHPEHNTLBNVRFPAGNAPCWRS	240
QY	241	RRTLAALAHCPYSKRVFLTHPATCYRTTCGPDSDPQCONGTCVCEGLDGYOCLCP	300
DB	241	RRTLAALAHCPYSKRVFLTHPATCYRTTCGPDSDPQCONGTCVCEGLDGYOCLCP	300
QY	241	RRTLAALAHCPYSKRVFLTHPATCYRTTCGPDSDPQCONGTCVCEGLDGYOCLCP	300
DB	241	RRTLAALAHCPYSKRVFLTHPATCYRTTCGPDSDPQCONGTCVCEGLDGYOCLCP	300
QY	301	LAFGGANCAKLSLECRVDLLFLDSSAGTTLDGLFRAKAVYKRPVRAVLSEDSRAVG	360
DB	301	LAFGGANCAKLSLECRVDLLFLDSSAGTTLDGLFRAKAVYKRPVRAVLSEDSRAVG	360
QY	301	LAFGGANCAKLSLECRVDLLFLDSSAGTTLDGLFRAKAVYKRPVRAVLSEDSRAVG	360
DB	301	LAFGGANCAKLSLECRVDLLFLDSSAGTTLDGLFRAKAVYKRPVRAVLSEDSRAVG	360
QY	361	VATYSRELLVAVPVGVGVDPDLVMSLDGIPRGGSTTLTGSALROAAREGSAATGGD	420
DB	361	VATYSRELLVAVPVGVGVDPDLVMSLDGIPRGGSTTLTGSALROAAREGSAATGGD	420
QY	361	VATYSRELLVAVPVGVGVDPDLVMSLDGIPRGGSTTLTGSALROAAREGSAATGGD	420
DB	361	VATYSRELLVAVPVGVGVDPDLVMSLDGIPRGGSTTLTGSALROAAREGSAATGGD	420

QY 421 RPRRVVVLTTSESHSEDEVAGPARHARARELLLLGVSEAVRALEETTSQPKVMYSDP 480
DB 421 RPRRVVVLTTSESHSEDEVAGPARHARARELLLLGVSEAVRALEETTSQPKVMYSDP 480
QY 481 QDLFNOIPELQGLKCSRQRCGRTOALDVFMLDTSASVGPENFAMQSFVRSALQFV 540
DB 481 QDLFNOIPELQGLKCSRQRCGRTOALDVFMLDTSASVGPENFAMQSFVRSALQFV 540
QY 541 NPDVTQGLVTVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVSGAGTALLHIYDKVMT 600
DB 541 NPDVTQGLVTVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVSGAGTALLHIYDKVMT 600
QY 601 VORGARPGVPAKVVVLTGSGAEDAAVPAQKLRNNGISVLYVGVGPVLSGLRRLAGPRD 660
DB 601 VORGARPGVPAKVVVLTGSGAEDAAVPAQKLRNNGISVLYVGVGPVLSGLRRLAGPRD 660
QY 661 SLIHVAAYADLRHODVLEMLCGEAKOPVNLCKPSCMNEGSCVYLQNGSYRCKCRDGE 720
DB 661 SLIHVAAYADLRHODVLEMLCGEAKOPVNLCKPSCMNEGSCVYLQNGSYRCKCRDGE 720
QY 721 GPHCENRFLRRP 732
DB 721 GPHCENRFLRRP 732

RESULT 2
US-10-274-177-3
; Sequence 3, Application US/10274177
; Publication No. US20040038225A1
; GENERAL INFORMATION:
; APPLICANT: MARKOWITZ, Sanford D.
; TITLE OF INVENTION: METHODS FOR CATEGORIZING PATIENTS
; FILE REFERENCE: CMRU-P01-003
; CURRENT APPLICATION NUMBER: US/10/274,177
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: US/10/229,345
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; LENGTH: 732
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-274-177-3

Query Match 100.0%; Score 3826; DB 4; Length 732;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITGKISAASKMMCSAAVDIMFLDGSNSVGKSFERSKHFAITVCDGLDI 60
DB 1 LOEVHVSKEITGKISAASKMMCSAAVDIMFLDGSNSVGKSFERSKHFAITVCDGLDI 60
QY 61 SPERRVAVGAFQSSSTPHLEPPLDSFSTQOEVKARIKRMVFKGRTTETLALKYLHRGUP 120
DB 61 SPERRVAVGAFQSSSTPHLEPPLDSFSTQOEVKARIKRMVFKGRTTETLALKYLHRGUP 120
QY 121 GGRNAAVPOILLIIVTDGKSQGDVALPSKQKKEGTVTVAVGVFRPMEELHALLASBPQ 180
DB 121 GGRNAAVPOILLIIVTDGKSQGDVALPSKQKKEGTVTVAVGVFRPMEELHALLASBPQ 180
QY 181 HVLAAQVEDATNGLFSTLSSAICSSATPDCRVEAHPCERHRTLEMBREPFAGNAPCWRGS 240
DB 181 HVLAAQVEDATNGLFSTLSSAICSSATPDCRVEAHPCERHRTLEMBREPFAGNAPCWRGS 240
QY 241 RRTLAVALAAHCPYSSMKRVFLTHPATCYRTTCGPCDSOPCQNGGTCEBGLDGYQCLP 300
DB 241 RRTLAVALAAHCPYSSMKRVFLTHPATCYRTTCGPCDSOPCQNGGTCEBGLDGYQCLP 300
QY 301 LAFGSGANCAKLSLECRVDLFLDSSAGTTLDGLFRAKVFYKRFVRAVLSBDSRAV 360
DB 301 LAFGSGANCAKLSLECRVDLFLDSSAGTTLDGLFRAKVFYKRFVRAVLSBDSRAV 360

QY 361 VATYSHELLVAAPVGEYQDVPDLWMSLDGI PRGGEPTLTGSAALRQAAERFGSATRTGQD 420
DB 361 VATYSHELLVAAPVGEYQDVPDLWMSLDGI PRGGEPTLTGSAALRQAAERFGSATRTGQD 420
QY 421 RPRRVVVLTTSESHSEDEVAGPARHARARELLLLGVSEAVRALEETTSQPKVMYSDP 480
DB 421 RPRRVVVLTTSESHSEDEVAGPARHARARELLLLGVSEAVRALEETTSQPKVMYSDP 480
QY 481 QDLFNOIPELQGLKCSRQRCGRTOALDVFMLDTSASVGPENFAMQSFVRSALQFV 540
DB 481 QDLFNOIPELQGLKCSRQRCGRTOALDVFMLDTSASVGPENFAMQSFVRSALQFV 540
QY 541 NPDVTQGLVTVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVSGAGTALLHIYDKVMT 600
DB 541 NPDVTQGLVTVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVSGAGTALLHIYDKVMT 600
QY 601 VORGARPGVPAKVVVLTGSGAEDAAVPAQKLRNNGISVLYVGVGPVLSGLRRLAGPRD 660
DB 601 VORGARPGVPAKVVVLTGSGAEDAAVPAQKLRNNGISVLYVGVGPVLSGLRRLAGPRD 660
QY 661 SLIHVAAYADLRHODVLEMLCGEAKOPVNLCKPSCMNEGSCVYLQNGSYRCKCRDGE 720
DB 661 SLIHVAAYADLRHODVLEMLCGEAKOPVNLCKPSCMNEGSCVYLQNGSYRCKCRDGE 720
QY 721 GPHCENRFLRRP 732
DB 721 GPHCENRFLRRP 732

RESULT 3
US-10-650-112-3
; Sequence 3, Application US/10650112
; Publication No. US20040110712A1
; GENERAL INFORMATION:
; APPLICANT: MARKOWITZ, Sanford D.
; TITLE OF INVENTION: METHODS FOR TREATING PATIENTS AND IDENTIFYING THERAPEUTICS
; FILE REFERENCE: CMRU-P01-044
; CURRENT APPLICATION NUMBER: US/10/650,112
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 10/274,177
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 10/229,245
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/406,296
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 732
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-650-112-3

Query Match 100.0%; Score 3826; DB 4; Length 732;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITGKISAASKMMCSAAVDIMFLDGSNSVGKSFERSKHFAITVCDGLDI 60
DB 1 LOEVHVSKEITGKISAASKMMCSAAVDIMFLDGSNSVGKSFERSKHFAITVCDGLDI 60
QY 61 SPERRVAVGAFQSSSTPHLEPPLDSFSTQOEVKARIKRMVFKGRTTETLALKYLHRGUP 120
DB 61 SPERRVAVGAFQSSSTPHLEPPLDSFSTQOEVKARIKRMVFKGRTTETLALKYLHRGUP 120
QY 121 GGRNAAVPOILLIIVTDGKSQGDVALPSKQKKEGTVTVAVGVFRPMEELHALLASBPQ 180
DB 121 GGRNAAVPOILLIIVTDGKSQGDVALPSKQKKEGTVTVAVGVFRPMEELHALLASBPQ 180
QY 181 HVLAAQVEDATNGLFSTLSSAICSSATPDCRVEAHPCERHRTLEMBREPFAGNAPCWRGS 240
DB 181 HVLAAQVEDATNGLFSTLSSAICSSATPDCRVEAHPCERHRTLEMBREPFAGNAPCWRGS 240

```

QY 241 RRTLAVLAHCPFYSMKRVFLTHPATCYRTTCPCDSDPCCONGGTCVBEGLDGYQCLCP 300
| | | | |
DB 241 RRTLAVLAHCPFYSMKRVFLTHPATCYRTTCPCDSDPCCONGGTCVBEGLDGYQCLCP 300
QY 301 LAFGBGANALKLSIECRVDLLFLDSSAGTTLDGFLRAKVFVKRFVRAVLSEDSBARVG 360
| | | | |
DB 301 LAFGBGANALKLSIECRVDLLFLDSSAGTTLDGFLRAKVFVKRFVRAVLSEDSBARVG 360
QY 361 VATTYRELLVAVPVGEYQDVPDLVMSLDGIPRGGPTLTGSALROAERGFSGATRTGDD 420
| | | | |
DB 361 VATTYRELLVAVPVGEYQDVPDLVMSLDGIPRGGPTLTGSALROAERGFSGATRTGDD 420
QY 421 RRRRVVLLTSHSEDEVAAGPARHARARELLLGVSBAVARALEEITGSPKXVMVYSDP 480
| | | | |
DB 421 RRRRVVLLTSHSEDEVAAGPARHARARELLLGVSBAVARALEEITGSPKXVMVYSDP 480
QY 481 ODLFNOIPELOKLCGRORPGCRTQALDLVFMLDTSASVGPENFAQMGSFVSCALQFEV 540
| | | | |
DB 481 ODLFNOIPELOKLCGRORPGCRTQALDLVFMLDTSASVGPENFAQMGSFVSCALQFEV 540
QY 541 NPDVTOVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAAYLGVSAGTALHITDKVMT 600
| | | | |
DB 541 NPDVTOVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAAYLGVSAGTALHITDKVMT 600
QY 601 VORGARPGVKAVVVLTVGSGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 660
| | | | |
DB 601 VORGARPGVKAVVVLTVGSGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 660
QY 661 SLIHVAAYADLRHYQDVLIEMLGCEBAKQPVNLCKPSPCMNEGSCVLYQNGSYRCKCKDWE 720
| | | | |
DB 661 SLIHVAAYADLRHYQDVLIEMLGCEBAKQPVNLCKPSPCMNEGSCVLYQNGSYRCKCKDWE 720
QY 721 GPHCENRFLRRP 732
| | | | |
DB 721 GPHCENRFLRRP 732

RESULT 4
US-10-649-591-3
; Sequence 3, Application US/10649591
; Publication No. US20060035237A1
; GENERAL INFORMATION:
; APPLICANT: Markowitz, Sanford D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; FILE REFERENCE: CMRU-P03-003
; CURRENT APPLICATION NUMBER: US/10/649,591
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 10/229,345
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: US 10/274,177
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 732
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-649-591-3

Query Match 100.0%; Score 3826; DB 5; Length 732;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 121 GGRNASVPQILITVTDKSGQDVALPSKQJKEGVTVFVAVGRFPRMEELHALASEPRGQ 180
| | | | |
QY 181 HTLAEQVEDATNGLPSTISSAICSSATPPDCRVNAHPCEHHTLTMVPEFAANAPCMWRS 240
| | | | |
DB 181 HTLAEQVEDATNGLPSTISSAICSSATPPDCRVNAHPCEHHTLTMVPEFAANAPCMWRS 240
QY 241 RRTLAVLAHCPFYSMKRVFLTHPATCYRTTCPCDSDPCCONGGTCVBEGLDGYQCLCP 300
| | | | |
DB 241 RRTLAVLAHCPFYSMKRVFLTHPATCYRTTCPCDSDPCCONGGTCVBEGLDGYQCLCP 300
QY 301 LAFGBGANALKLSIECRVDLLFLDSSAGTTLDGFLRAKVFVKRFVRAVLSEDSBARVG 360
| | | | |
DB 301 LAFGBGANALKLSIECRVDLLFLDSSAGTTLDGFLRAKVFVKRFVRAVLSEDSBARVG 360
QY 361 VATTYRELLVAVPVGEYQDVPDLVMSLDGIPRGGPTLTGSALROAERGFSGATRTGDD 420
| | | | |
DB 361 VATTYRELLVAVPVGEYQDVPDLVMSLDGIPRGGPTLTGSALROAERGFSGATRTGDD 420
QY 421 RRRRVVLLTSHSEDEVAAGPARHARARELLLGVSBAVARALEEITGSPKXVMVYSDP 480
| | | | |
DB 421 RRRRVVLLTSHSEDEVAAGPARHARARELLLGVSBAVARALEEITGSPKXVMVYSDP 480
QY 481 ODLFNOIPELOKLCGRORPGCRTQALDLVFMLDTSASVGPENFAQMGSFVSCALQFEV 540
| | | | |
DB 481 ODLFNOIPELOKLCGRORPGCRTQALDLVFMLDTSASVGPENFAQMGSFVSCALQFEV 540
QY 541 NPDVTOVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAAYLGVSAGTALHITDKVMT 600
| | | | |
DB 541 NPDVTOVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAAYLGVSAGTALHITDKVMT 600
QY 601 VORGARPGVKAVVVLTVGSGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 660
| | | | |
DB 601 VORGARPGVKAVVVLTVGSGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 660
QY 661 SLIHVAAYADLRHYQDVLIEMLGCEBAKQPVNLCKPSPCMNEGSCVLYQNGSYRCKCKDWE 720
| | | | |
DB 661 SLIHVAAYADLRHYQDVLIEMLGCEBAKQPVNLCKPSPCMNEGSCVLYQNGSYRCKCKDWE 720
QY 721 GPHCENRFLRRP 732
| | | | |
DB 721 GPHCENRFLRRP 732

RESULT 5
US-11-041-788-3
; Sequence 3, Application US/11041788
; Publication No. US2005023353A1
; GENERAL INFORMATION:
; APPLICANT: Markowitz, Sanford D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; FILE REFERENCE: CMRU-P04-003
; CURRENT APPLICATION NUMBER: US/11/041,788
; PRIOR FILING DATE: 2005-01-21
; PRIOR APPLICATION NUMBER: US 10/649,591
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 10/274,177
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: US 10/229,345
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: US 10/087,080
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 60/284,555
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 60/281,149
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/272,206
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 732

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-041-788-3

Query Match      100.0%; Score 3826; DB 6; Length 732;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNVGKSFERSKHFAITVCDGLDI 60
DB 1 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNVGKSFERSKHFAITVCDGLDI 60
QY 61 SPERVVGAFOFSSSTPHLEFPDLSFSTQGEVKARIKRVFKGRTETELAKYLLHRGLP 120
DB 61 SPERVVGAFOFSSSTPHLEFPDLSFSTQGEVKARIKRVFKGRTETELAKYLLHRGLP 120
QY 121 GGRNASVPOILLIIVTDGKSQGDVALPSKQLKRGVTVFAVGVPFPMWEELHALASBPQQ 180
DB 121 GGRNASVPOILLIIVTDGKSQGDVALPSKQLKRGVTVFAVGVPFPMWEELHALASBPQQ 180
QY 181 HVLAEQVEDATNGIFSTLSSSAICSSATPDCEVAHPCERHRTLEMYREPAAGNAPCMRGS 240
DB 181 HVLAEQVEDATNGIFSTLSSSAICSSATPDCEVAHPCERHRTLEMYREPAAGNAPCMRGS 240
QY 241 RRTLAIVLAHCPFSYKRVFLTHPATCYRTTCGPGCDSPQONGSTCVBGLDGYQCLCP 300
DB 241 RRTLAIVLAHCPFSYKRVFLTHPATCYRTTCGPGCDSPQONGSTCVBGLDGYQCLCP 300
QY 301 LAFGEANCAKLSLECRVDLLFLDSSAGTTLDGFLRAVFKRVFRAVLSSEDSRAVG 360
DB 301 LAFGEANCAKLSLECRVDLLFLDSSAGTTLDGFLRAVFKRVFRAVLSSEDSRAVG 360
QY 361 VATYSRELLVAVPGEYQDVPDLVWSLDGIPFRGGPTLGSALRQAERGFSGATRTGOD 420
DB 361 VATYSRELLVAVPGEYQDVPDLVWSLDGIPFRGGPTLGSALRQAERGFSGATRTGOD 420
QY 421 RPRRVVLLTSHSEDEBAGPARHARARELLLLGVSEAVRALEIEITGSPKHWVWVSDP 480
DB 421 RPRRVVLLTSHSEDEBAGPARHARARELLLLGVSEAVRALEIEITGSPKHWVWVSDP 480
QY 481 ODLFNOIPELOGLCSROPGRCTOALDVFMLDTSASVGEPMFAMQSPVSCALQEPV 540
DB 481 ODLFNOIPELOGLCSROPGRCTOALDVFMLDTSASVGEPMFAMQSPVSCALQEPV 540
QY 541 NPDVTOVGLVYVSGQVTAFGDLTKPTRAMLRALISQAPYLGVSAGTALLHIYDKWT 600
DB 541 NPDVTOVGLVYVSGQVTAFGDLTKPTRAMLRALISQAPYLGVSAGTALLHIYDKWT 600
QY 601 VORGAAPGVPAKAVVLTGGRGAEDAAVPAOKLRNNGISVLVVGVPVLSBGLRRLAGPRD 660
DB 601 VORGAAPGVPAKAVVLTGGRGAEDAAVPAOKLRNNGISVLVVGVPVLSBGLRRLAGPRD 660
QY 661 SLIHVAAYADLRHODVLEMLGGEAKOPVNLCKPSPCMNEGSCVLONSYRCKCKDME 720
DB 661 SLIHVAAYADLRHODVLEMLGGEAKOPVNLCKPSPCMNEGSCVLONSYRCKCKDME 720
QY 721 GPHCENRFLRRP 732
DB 721 GPHCENRFLRRP 732

RESULT 6
US-10-295-027-124
; Sequence 124, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Ataz, Daniel
; APPLICANT: Atiz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glyne, Richard
; APPLICANT: Hevel, Peter A.
; APPLICANT: Mack, David H.

```

```

; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 124
; LENGTH: 755
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-124

Query Match      100.0%; Score 3826; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNVGKSFERSKHFAITVCDGLDI 60
DB 24 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNVGKSFERSKHFAITVCDGLDI 83
QY 61 SPERVVGAFOFSSSTPHLEFPDLSFSTQGEVKARIKRVFKGRTETELAKYLLHRGLP 120
DB 84 SPERVVGAFOFSSSTPHLEFPDLSFSTQGEVKARIKRVFKGRTETELAKYLLHRGLP 143
QY 121 GGRNASVPOILLIIVTDGKSQGDVALPSKQLKRGVTVFAVGVPFPMWEELHALASBPQQ 180
DB 144 GGRNASVPOILLIIVTDGKSQGDVALPSKQLKRGVTVFAVGVPFPMWEELHALASBPQQ 203
QY 181 HVLAEQVEDATNGIFSTLSSSAICSSATPDCEVAHPCERHRTLEMYREPAAGNAPCMRGS 240
DB 204 HVLAEQVEDATNGIFSTLSSSAICSSATPDCEVAHPCERHRTLEMYREPAAGNAPCMRGS 263
QY 241 RRTLAIVLAHCPFSYKRVFLTHPATCYRTTCGPGCDSPQONGSTCVBGLDGYQCLCP 300
DB 264 RRTLAIVLAHCPFSYKRVFLTHPATCYRTTCGPGCDSPQONGSTCVBGLDGYQCLCP 323
QY 301 LAFGEANCAKLSLECRVDLLFLDSSAGTTLDGFLRAVFKRVFRAVLSSEDSRAVG 360
DB 324 LAFGEANCAKLSLECRVDLLFLDSSAGTTLDGFLRAVFKRVFRAVLSSEDSRAVG 383
QY 361 VATYSRELLVAVPGEYQDVPDLVWSLDGIPFRGGPTLGSALRQAERGFSGATRTGOD 420
DB 384 VATYSRELLVAVPGEYQDVPDLVWSLDGIPFRGGPTLGSALRQAERGFSGATRTGOD 443
QY 421 RPRRVVLLTSHSEDEBAGPARHARARELLLLGVSEAVRALEIEITGSPKHWVWVSDP 480
DB 444 RPRRVVLLTSHSEDEBAGPARHARARELLLLGVSEAVRALEIEITGSPKHWVWVSDP 503

```

QY 481 ODLENOIPELOKCLSGRORPGCRTOALDVFMLDTSASVGPENFAMOMOSFYVSCALQFEV 540
DB 504 ODLENOIPELOKCLSGRORPGCRTOALDVFMLDTSASVGPENFAMOMOSFYVSCALQFEV 563
QY 541 NPDVTQVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKVMT 600
DB 564 NPDVTQVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKVMT 623
QY 601 VORGARPGVPKAVVVLITGGRGAEDAAVPAQKLRNNGISYLVVGVGVLSBGLRRLAGPRD 660
DB 624 VORGARPGVPKAVVVLITGGRGAEDAAVPAQKLRNNGISYLVVGVGVLSBGLRRLAGPRD 683
QY 661 SLIHYAAVADLRHYODVLEIWLCEAKOPVNLCKPSPCMNEGSCVYLQNGSYRCKCKDGM 720
DB 684 SLIHYAAVADLRHYODVLEIWLCEAKOPVNLCKPSPCMNEGSCVYLQNGSYRCKCKDGM 743
QY 721 GPHCENRFLRRP 732
DB 744 GPHCENRFLRRP 755

RESULT 7
US-10-173-999-87
; Sequence 87, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 755
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-87

Query Match 100.0%; Score 3826; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITGKISAASKMMCSAAVDIMFLDGSNSVKGSPERSKPAITVCGGLD 60
DB 24 LOEVHVSKEITGKISAASKMMCSAAVDIMFLDGSNSVKGSPERSKPAITVCGGLD 83
QY 61 SPERVVGAFOFSTPHLEPPLDSFSTQOEVKARIKMWPKGRTETELAKTLLHRLGP 120
DB 84 SPERVVGAFOFSTPHLEPPLDSFSTQOEVKARIKMWPKGRTETELAKTLLHRLGP 143
QY 121 GGNASVPOILLIIVTDGKSGDVALPSKQKRGVTVFAVGVPFPMWELHALASEBRQ 180
DB 144 GGNASVPOILLIIVTDGKSGDVALPSKQKRGVTVFAVGVPFPMWELHALASEBRQ 203
QY 181 HTLLAQVEDATNGLFSTLSSSAICSSATPDCEVBAHPCEHRTLDMVREPAAGNAPCMRG 240
DB 204 HTLLAQVEDATNGLFSTLSSSAICSSATPDCEVBAHPCEHRTLDMVREPAAGNAPCMRG 263
QY 241 RRTLAVALAHCPFSMWKRVFLTHPATCYRTTCPCDOPCONGTCVPEGLDGYOCLCP 300

DB 264 RRTLAVALAHCPFSMWKRVFLTHPATCYRTTCPCDOPCONGTCVPEGLDGYOCLCP 323
QY 301 LAFGEANCAKLSIECRVDLFLDSSAGTTLDGFLRAKVPFKFVRAVLSBDSRARVG 360
DB 324 LAFGEANCAKLSIECRVDLFLDSSAGTTLDGFLRAKVPFKFVRAVLSBDSRARVG 383
QY 361 VATYRELLVANVPGEYQDPDLVWSLGGIIPRGGFTLTGSLALROAERGFSSATRTGQD 420
DB 384 VATYRELLVANVPGEYQDPDLVWSLGGIIPRGGFTLTGSLALROAERGFSSATRTGQD 443
QY 421 RPRRVVLLTSHSSEDEVAAPARHARARELLLGVSSEVARELEITGSPHVMVYSP 480
DB 444 RPRRVVLLTSHSSEDEVAAPARHARARELLLGVSSEVARELEITGSPHVMVYSP 503
QY 481 ODLENOIPELOKCLSGRORPGCRTOALDVFMLDTSASVGPENFAMOMOSFYVSCALQFEV 540
DB 504 ODLENOIPELOKCLSGRORPGCRTOALDVFMLDTSASVGPENFAMOMOSFYVSCALQFEV 563
QY 541 NPDVTQVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKVMT 600
DB 564 NPDVTQVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALLHIYDKVMT 623
QY 601 VORGARPGVPKAVVVLITGGRGAEDAAVPAQKLRNNGISYLVVGVGVLSBGLRRLAGPRD 660
DB 624 VORGARPGVPKAVVVLITGGRGAEDAAVPAQKLRNNGISYLVVGVGVLSBGLRRLAGPRD 683
QY 661 SLIHYAAVADLRHYODVLEIWLCEAKOPVNLCKPSPCMNEGSCVYLQNGSYRCKCKDGM 720
DB 684 SLIHYAAVADLRHYODVLEIWLCEAKOPVNLCKPSPCMNEGSCVYLQNGSYRCKCKDGM 743
QY 721 GPHCENRFLRRP 732
DB 744 GPHCENRFLRRP 755

RESULT 8
US-10-229-345-14
; Sequence 14, Application US/10229345
; Publication No. US20040038220A1
; GENERAL INFORMATION:
; APPLICANT: MARKOWITZ, Sanford D.
; TITLE OF INVENTION: METHODS FOR CATEGORIZING PATIENTS
; FILE REFERENCE: CMU-P01-003
; CURRENT APPLICATION NUMBER: US/10/229,345
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 755
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-229-345-14

Query Match 100.0%; Score 3826; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITGKISAASKMMCSAAVDIMFLDGSNSVKGSPERSKPAITVCGGLD 60
DB 24 LOEVHVSKEITGKISAASKMMCSAAVDIMFLDGSNSVKGSPERSKPAITVCGGLD 83
QY 61 SPERVVGAFOFSTPHLEPPLDSFSTQOEVKARIKMWPKGRTETELAKTLLHRLGP 120
DB 84 SPERVVGAFOFSTPHLEPPLDSFSTQOEVKARIKMWPKGRTETELAKTLLHRLGP 143
QY 121 GGNASVPOILLIIVTDGKSGDVALPSKQKRGVTVFAVGVPFPMWELHALASEBRQ 180
DB 144 GGNASVPOILLIIVTDGKSGDVALPSKQKRGVTVFAVGVPFPMWELHALASEBRQ 203
QY 181 HTLLAQVEDATNGLFSTLSSSAICSSATPDCEVBAHPCEHRTLDMVREPAAGNAPCMRG 240
DB 204 HTLLAQVEDATNGLFSTLSSSAICSSATPDCEVBAHPCEHRTLDMVREPAAGNAPCMRG 263

QY 241 RRTLAVALAHCPRYSWKRVFLTHPATCYRTTCPCPCDSOPCONGTCTVPEGLDGYOCLCP 300
DB 264 RRTLAVALAHCPRYSWKRVFLTHPATCYRTTCPCPCDSOPCONGTCTVPEGLDGYOCLCP 323
QY 301 LAFGEANCALKLSLECRVDLFLDSSAGTTLDFLRKAVFYKRFVRAVLSBDSBARVG 360
DB 324 LAFGEANCALKLSLECRVDLFLDSSAGTTLDFLRKAVFYKRFVRAVLSBDSBARVG 383
QY 361 VATYSRELLVAVPVEGYODPDLVWSLDGI PRGGPTLTGSALRQAERGFSGATRTGOD 420
DB 384 VATYSRELLVAVPVEGYODPDLVWSLDGI PRGGPTLTGSALRQAERGFSGATRTGOD 443
QY 421 RPRRVVVLTTSHSEDEVA GPARRARARELLLGVSSEAVRALEBITGSPKXVMVYSDP 480
DB 444 RPRRVVVLTTSHSEDEVA GPARRARARELLLGVSSEAVRALEBITGSPKXVMVYSDP 503
QY 481 ODLENOIPELQKLCGRORPGCRTOALDVFMLDTSASVGPENPAOMOSFVRSALQFEV 540
DB 504 ODLENOIPELQKLCGRORPGCRTOALDVFMLDTSASVGPENPAOMOSFVRSALQFEV 563
QY 541 NPDVTQGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALHIYDKVMT 600
DB 564 NPDVTQGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALHIYDKVMT 623
QY 601 VORGARPGVPKAVVYLTGGRGAEDAAVPAOKLRNNGISVLYVGVGPVLSSEGLRRLAGPRD 660
DB 624 VORGARPGVPKAVVYLTGGRGAEDAAVPAOKLRNNGISVLYVGVGPVLSSEGLRRLAGPRD 683
QY 661 SLIHVAAYADLRHYODVLIEMLCGEAKOPVNLCKPSCMNEGSCVLONGSYRCKCRDGE 720
DB 684 SLIHVAAYADLRHYODVLIEMLCGEAKOPVNLCKPSCMNEGSCVLONGSYRCKCRDGE 743
QY 721 GPHCENRFLRRP 732
DB 744 GPHCENRFLRRP 755

RESULT 9
US-10-274-177-14
; Sequence 14, Application US/10274177
; Publication No. US20040038225A1
; GENERAL INFORMATION:
; APPLICANT: MARKOWITZ, Sanford D.
; TITLE OF INVENTION: METHODS FOR CATEGORIZING PATIENTS
; FILE REFERENCE: CMRU-P01-003
; CURRENT APPLICATION NUMBER: US/10/274,177
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: US/10/229,345
; PRIOR FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 755
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-274-177-14

Query Match 100.0%; Score 3826; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITGKISAAKMMWCSAAVDIMFLDGSNVGKSPERSKHFAITVCDGLDI 60
DB 24 LOEVHVSKEITGKISAAKMMWCSAAVDIMFLDGSNVGKSPERSKHFAITVCDGLDI 83

QY 61 SBERVAVGAFOSSPTPHLEFPDLSFSTOEVKARIRKMPKGRTEITELATYLLHRGIP 120
DB 84 SBERVAVGAFOSSPTPHLEFPDLSFSTOEVKARIRKMPKGRTEITELATYLLHRGIP 143

QY 121 GGRNAAVPOILLIIVTDGSKQGVVALPSKOLKRGVTVFAVGVRFPRWEELHALASPRQO 180
DB 144 GGRNAAVPOILLIIVTDGSKQGVVALPSKOLKRGVTVFAVGVRFPRWEELHALASPRQO 203

QY 181 HYLLAEQVEDATNGLFSTLLSSAICSSATPDRCVREAHPCENHRTLBMVREFPAGNAPCMRGS 240
DB 204 HYLLAEQVEDATNGLFSTLLSSAICSSATPDRCVREAHPCENHRTLBMVREFPAGNAPCMRGS 263
QY 241 RRTLAVALAHCPRYSWKRVFLTHPATCYRTTCPCPCDSOPCONGTCTVPEGLDGYOCLCP 300
DB 264 RRTLAVALAHCPRYSWKRVFLTHPATCYRTTCPCPCDSOPCONGTCTVPEGLDGYOCLCP 323
QY 301 LAFGEANCALKLSLECRVDLFLDSSAGTTLDFLRKAVFYKRFVRAVLSBDSBARVG 360
DB 324 LAFGEANCALKLSLECRVDLFLDSSAGTTLDFLRKAVFYKRFVRAVLSBDSBARVG 383
QY 361 VATYSRELLVAVPVEGYODPDLVWSLDGI PRGGPTLTGSALRQAERGFSGATRTGOD 420
DB 384 VATYSRELLVAVPVEGYODPDLVWSLDGI PRGGPTLTGSALRQAERGFSGATRTGOD 443
QY 421 RPRRVVVLTTSHSEDEVA GPARRARARELLLGVSSEAVRALEBITGSPKXVMVYSDP 480
DB 444 RPRRVVVLTTSHSEDEVA GPARRARARELLLGVSSEAVRALEBITGSPKXVMVYSDP 503
QY 481 ODLENOIPELQKLCGRORPGCRTOALDVFMLDTSASVGPENPAOMOSFVRSALQFEV 540
DB 504 ODLENOIPELQKLCGRORPGCRTOALDVFMLDTSASVGPENPAOMOSFVRSALQFEV 563
QY 541 NPDVTQGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALHIYDKVMT 600
DB 564 NPDVTQGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGVSAGTALHIYDKVMT 623
QY 601 VORGARPGVPKAVVYLTGGRGAEDAAVPAOKLRNNGISVLYVGVGPVLSSEGLRRLAGPRD 660
DB 624 VORGARPGVPKAVVYLTGGRGAEDAAVPAOKLRNNGISVLYVGVGPVLSSEGLRRLAGPRD 683
QY 661 SLIHVAAYADLRHYODVLIEMLCGEAKOPVNLCKPSCMNEGSCVLONGSYRCKCRDGE 720
DB 684 SLIHVAAYADLRHYODVLIEMLCGEAKOPVNLCKPSCMNEGSCVLONGSYRCKCRDGE 743
QY 721 GPHCENRFLRRP 732
DB 744 GPHCENRFLRRP 755

RESULT 10
US-10-650-112-14
; Sequence 14, Application US/10650112
; Publication No. US2004010712A1
; GENERAL INFORMATION:
; APPLICANT: MARKOWITZ, Sanford D.
; TITLE OF INVENTION: METHODS FOR TREATING PATIENTS AND IDENTIFYING THERAPEUTICS
; FILE REFERENCE: CMRU-P01-044
; CURRENT APPLICATION NUMBER: US/10/650,112
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 10/274,177
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 10/229,245
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/406,296
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 755
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-650-112-14

Query Match 100.0%; Score 3826; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITGKISAAKMMWCSAAVDIMFLDGSNVGKSPERSKHFAITVCDGLDI 60
DB 24 LOEVHVSKEITGKISAAKMMWCSAAVDIMFLDGSNVGKSPERSKHFAITVCDGLDI 83


```

Qy 61 SPBRVGAFOFSSPTHEPLDPSFSTQOEVKARIKRMVFGKGRTELTALKYLHRGP 120
| 121 GGRNMSVPOILIIITVDGKSQGVNLPKQKRGVTVFAVGVFPMEELHALASPRQ 180
| 144 GGRNMSVPOILIIITVDGKSQGVNLPKQKRGVTVFAVGVFPMEELHALASPRQ 203
| 181 HYLAEQVEDATNGLFSTLSSAICSSATPDCRVEAHPCHEHRTLEKVRFPAGNAPCWRGS 240
| 204 HYLAEQVEDATNGLFSTLSSAICSSATPDCRVEAHPCHEHRTLEKVRFPAGNAPCWRGS 263
| 241 RRTLAVALAHCPYSKRVFLTHPATCYRTTCGPGCDSPCONGGTCVPEGLDGYOCLCP 300
| 264 RRTLAVALAHCPYSKRVFLTHPATCYRTTCGPGCDSPCONGGTCVPEGLDGYOCLCP 323
| 301 LAFGEANCALKLSLECRVDLFLDSSAGTTLDFLRKAVFKRVFRAVLSDSBARVG 360
| 324 LAFGEANCALKLSLECRVDLFLDSSAGTTLDFLRKAVFKRVFRAVLSDSBARVG 383
| 361 VATYSRELLVAVPVGEYQVDPDLVMSLDGI PRFGGPTLTGSALRQAEFGSATRTGOD 420
| 384 VATYSRELLVAVPVGEYQVDPDLVMSLDGI PRFGGPTLTGSALRQAEFGSATRTGOD 443
| 421 RPRRVVVLTTESHSEDEVGAPARHARARELLLVGSSEAVRALEBITSPKHMVWYSDP 480
| 444 RPRRVVVLTTESHSEDEVGAPARHARARELLLVGSSEAVRALEBITSPKHMVWYSDP 503
| 481 ODLENOIPELOGKLCRQRPGRCTQALDLVFMLDTSASVGPENFAMQSFVNSCALQFV 540
| 504 ODLENOIPELOGKLCRQRPGRCTQALDLVFMLDTSASVGPENFAMQSFVNSCALQFV 563
| 541 NPDVTQVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVGSAGTALLHIYDKWMT 600
| 564 NPDVTQVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVGSAGTALLHIYDKWMT 623
| 601 VORGARPGVPAVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSGLRRLAGPRD 660
| 624 VORGARPGVPAVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSGLRRLAGPRD 683
| 661 SLIHVAAYADLRHODVLEMLCGEAKOPVNLCKPSPCMMEGSCVLYQNGSYRCKCDGME 720
| 684 SLIHVAAYADLRHODVLEMLCGEAKOPVNLCKPSPCMMEGSCVLYQNGSYRCKCDGME 743
|
Qy 721 GPHCENRFLRRP 732
| 744 GPHCENRFLRRP 755
|
Db

```

RESULT 11

```

; US-10-649-591-14
; Sequence 14, Application US/10649591
; Publication No. US20060035237A1
; GENERAL INFORMATION:
; APPLICANT: Markowitz, Sanford D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; FILE REFERENCE: CMRU-P03-003
; CURRENT APPLICATION NUMBER: US/10/649,591
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 10/229,345
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: US 10/274,177
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 755
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-649-591-14

```

Query Match

100.0%; Score 3826; DB 5; Length 755;

```

Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LOEVHVSKEITIGKISAAKMMKCSAAVDMFLDGSNSVSKGSFFESKFAITVCGLDI 60
| 24 LOEVHVSKEITIGKISAAKMMKCSAAVDMFLDGSNSVSKGSFFESKFAITVCGLDI 83
| 61 SPBRVGAFOFSSPTHEPLDPSFSTQOEVKARIKRMVFGKGRTELTALKYLHRGP 120
| 84 SPBRVGAFOFSSPTHEPLDPSFSTQOEVKARIKRMVFGKGRTELTALKYLHRGP 143
| 121 GGRNMSVPOILIIITVDGKSQGVNLPKQKRGVTVFAVGVFPMEELHALASPRQ 180
| 144 GGRNMSVPOILIIITVDGKSQGVNLPKQKRGVTVFAVGVFPMEELHALASPRQ 203
| 181 HYLAEQVEDATNGLFSTLSSAICSSATPDCRVEAHPCHEHRTLEKVRFPAGNAPCWRGS 240
| 204 HYLAEQVEDATNGLFSTLSSAICSSATPDCRVEAHPCHEHRTLEKVRFPAGNAPCWRGS 263
| 241 RRTLAVALAHCPYSKRVFLTHPATCYRTTCGPGCDSPCONGGTCVPEGLDGYOCLCP 300
| 264 RRTLAVALAHCPYSKRVFLTHPATCYRTTCGPGCDSPCONGGTCVPEGLDGYOCLCP 323
| 301 LAFGEANCALKLSLECRVDLFLDSSAGTTLDFLRKAVFKRVFRAVLSDSBARVG 360
| 324 LAFGEANCALKLSLECRVDLFLDSSAGTTLDFLRKAVFKRVFRAVLSDSBARVG 383
| 361 VATYSRELLVAVPVGEYQVDPDLVMSLDGI PRFGGPTLTGSALRQAEFGSATRTGOD 420
| 384 VATYSRELLVAVPVGEYQVDPDLVMSLDGI PRFGGPTLTGSALRQAEFGSATRTGOD 443
| 421 RPRRVVVLTTESHSEDEVGAPARHARARELLLVGSSEAVRALEBITSPKHMVWYSDP 480
| 444 RPRRVVVLTTESHSEDEVGAPARHARARELLLVGSSEAVRALEBITSPKHMVWYSDP 503
| 481 ODLENOIPELOGKLCRQRPGRCTQALDLVFMLDTSASVGPENFAMQSFVNSCALQFV 540
| 504 ODLENOIPELOGKLCRQRPGRCTQALDLVFMLDTSASVGPENFAMQSFVNSCALQFV 563
| 541 NPDVTQVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVGSAGTALLHIYDKWMT 600
| 564 NPDVTQVGLVYVGSQVQTAFGDLTKPTRAAMLRAISQAPYLGGVGSAGTALLHIYDKWMT 623
| 601 VORGARPGVPAVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSGLRRLAGPRD 660
| 624 VORGARPGVPAVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSGLRRLAGPRD 683
| 661 SLIHVAAYADLRHODVLEMLCGEAKOPVNLCKPSPCMMEGSCVLYQNGSYRCKCDGME 720
| 684 SLIHVAAYADLRHODVLEMLCGEAKOPVNLCKPSPCMMEGSCVLYQNGSYRCKCDGME 743
|
Qy 721 GPHCENRFLRRP 732
| 744 GPHCENRFLRRP 755
|
Db

```

RESULT 12

```

; US-11-041-788-14
; Sequence 14, Application US/11041788
; Publication No. US2005023353A1
; GENERAL INFORMATION:
; APPLICANT: Markowitz, Sanford D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; FILE REFERENCE: CMRU-P04-003
; CURRENT APPLICATION NUMBER: US/11/041,788
; PRIOR FILING DATE: 2005-01-21
; PRIOR APPLICATION NUMBER: US 10/649,591
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: US 10/274,177
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: US 10/229,345
; PRIOR FILING DATE: 2002-08-26

```

```

; PRIOR APPLICATION NUMBER: US 10/087,080
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 60/284,555
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: US 60/281,149
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/272,206
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 755
; TYPE: PR1
; ORGANISM: Homo sapiens
US-11-041-788-14

```

```

Query Match      100.0%; Score 3826; DB 6; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LOEVHVSKEITGKISAAKMMWCSAAVDIMFLDGSNSVGKSFERSKHFATVCDGLDI 60
DB 24 LOEVHVSKEITGKISAAKMMWCSAAVDIMFLDGSNSVGKSFERSKHFATVCDGLDI 83
QY 61 SPERRVGAFOFSSSTPHLEFPDLSFSTQOEVRARIRKMFKGGRTETELAKYLHRGLP 120
DB 84 SPERRVGAFOFSSSTPHLEFPDLSFSTQOEVRARIRKMFKGGRTETELAKYLHRGLP 143
QY 121 GGRNAAVPOILLIIVTDGKSQGDVALPSKQKRGVTVFAVGVFRPMBEHLALASBPQ 180
DB 144 GGRNAAVPOILLIIVTDGKSQGDVALPSKQKRGVTVFAVGVFRPMBEHLALASBPQ 203
QY 181 HVLAAQVEDATNGLFSTLSSAICSSATPDCVEAHPCHEHRTLEWREBAGNAPCMRGS 240
DB 204 HVLAAQVEDATNGLFSTLSSAICSSATPDCVEAHPCHEHRTLEWREBAGNAPCMRGS 263
QY 241 RRTLAIVLAHCPFYSMKRVFLTHPATCYRTTCGPGCDSPQCGNGTCVPEGLDGYOCLCP 300
DB 264 RRTLAIVLAHCPFYSMKRVFLTHPATCYRTTCGPGCDSPQCGNGTCVPEGLDGYOCLCP 323
QY 301 LAATGEANCAKLSLECRVDLFLDSSAGTTLDGFLRAVFAVYKRFVRAVLSDSBARVG 360
DB 324 LAATGEANCAKLSLECRVDLFLDSSAGTTLDGFLRAVFAVYKRFVRAVLSDSBARVG 383
QY 361 VATYSRELLVAVGEYQVDPDLVWSLDGIPFRGPTLTGSALRQAABRGFSATRTGQD 420
DB 384 VATYSRELLVAVGEYQVDPDLVWSLDGIPFRGPTLTGSALRQAABRGFSATRTGQD 443
QY 421 RPRRVVVLTTESSEBDEBAGPAAHARARELLLVGSEAVRALEBITGSPKRVMYSDP 480
DB 444 RPRRVVVLTTESSEBDEBAGPAAHARARELLLVGSEAVRALEBITGSPKRVMYSDP 503
QY 481 ODLFNQIPELQGLCSRRPGGRTQALDVFMLDTSASVGEFPAQMOSVRSCLQFXY 540
DB 504 ODLFNQIPELQGLCSRRPGGRTQALDVFMLDTSASVGEFPAQMOSVRSCLQFXY 563
QY 541 NPDVTQGLVVGSGQVTAAGLDTKPTRAAMLRALISQAPYLGVSAGTALLHIYDKWMT 600
DB 564 NPDVTQGLVVGSGQVTAAGLDTKPTRAAMLRALISQAPYLGVSAGTALLHIYDKWMT 623
QY 601 VORGAAPGVKAVVLTGGRGAEDAAVPAQKLRNNGISVLVGVGVLSBGLRRLAEPD 660
DB 624 VORGAAPGVKAVVLTGGRGAEDAAVPAQKLRNNGISVLVGVGVLSBGLRRLAEPD 683
QY 661 SLTHVAAYADLRHNOVLTEMLGSEAKOPVNLCKPSCMNEGSCVLONGSYRCKCKDME 720
DB 684 SLTHVAAYADLRHNOVLTEMLGSEAKOPVNLCKPSCMNEGSCVLONGSYRCKCKDME 743
QY 721 GPHCENRFLRRP 732
DB 744 GPHCENRFLRRP 755

```

```

RESULT 13
US-10-295-027-1186
; Sequence 1186, Application US/10295027
; Publication No. US20030232350A1
GENERAL INFORMATION:

```

```

; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natsaba
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1186
; LENGTH: 755
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-295-027-1186

```

```

Query Match      99.9%; Score 3822; DB 4; Length 755;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 731; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LOEVHVSKEITGKISAAKMMWCSAAVDIMFLDGSNSVGKSFERSKHFATVCDGLDI 60
DB 24 LOEVHVSKEITGKISAAKMMWCSAAVDIMFLDGSNSVGKSFERSKHFATVCDGLDI 83
QY 61 SPERRVGAFOFSSSTPHLEFPDLSFSTQOEVRARIRKMFKGGRTETELAKYLHRGLP 120
DB 84 SPERRVGAFOFSSSTPHLEFPDLSFSTQOEVRARIRKMFKGGRTETELAKYLHRGLP 143
QY 121 GGRNAAVPOILLIIVTDGKSQGDVALPSKQKRGVTVFAVGVFRPMBEHLALASBPQ 180
DB 144 GGRNAAVPOILLIIVTDGKSQGDVALPSKQKRGVTVFAVGVFRPMBEHLALASBPQ 203
QY 181 HVLAAQVEDATNGLFSTLSSAICSSATPDCVEAHPCHEHRTLEWREBAGNAPCMRGS 240
DB 204 HVLAAQVEDATNGLFSTLSSAICSSATPDCVEAHPCHEHRTLEWREBAGNAPCMRGS 263
QY 241 RRTLAIVLAHCPFYSMKRVFLTHPATCYRTTCGPGCDSPQCGNGTCVPEGLDGYOCLCP 300
DB 264 RRTLAIVLAHCPFYSMKRVFLTHPATCYRTTCGPGCDSPQCGNGTCVPEGLDGYOCLCP 323

```


QY 301 LAFGEANCALKLSLECRVDLFLDSSAGTTLDGFLRAKVFYKRFVRAVLSEDSBARVG 360
 DB 324 LAFGEANCALKLSLECRVDLFLDSSAGTTLDGFLRAKVFYKRFVRAVLSEDSBARVG 383
 QY 361 VATTYRELLVAAPVGEYQDPDLVMSLDGI PRGPGPTLTGSALRQAERGFSGATRTGDD 420
 DB 384 VATTYRELLVAAPVGEYQDPDLVMSLDGI PRGPGPTLTGSALRQAERGFSGATRTGDD 443
 QY 421 RRRRVVLLTSHSEDEBVAQPARHARARELLLGVSSEAVRALEBETITGSPKRWVYSDP 480
 DB 444 RRRRVVLLTSHSEDEBVAQPARHARARELLLGVSSEAVRALEBETITGSPKRWVYSDP 503
 QY 481 ODLENOIPELOGLCSRGPRGCRTOALDVFMLDTSASVGPENFAQMOGFSVSCALQFEV 540
 DB 504 ODLENOIPELOGLCSRGPRGCRTOALDVFMLDTSASVGPENFAQMOGFSVSCALQFEV 563
 QY 541 NPDVTQGLVYVGSQVQTAFGDTPKTRAMLRASQAAPYLGVSAGTALLHIYDKWNT 600
 DB 564 NPDVTQGLVYVGSQVQTAFGDTPKTRAMLRASQAAPYLGVSAGTALLHIYDKWNT 623
 QY 601 VORGARPGVPKAVVYLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 660
 DB 624 VORGARPGVPKAVVYLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHYHODVLEMLCGEAKOPVNLCKSPCMNBGSCTVLQNGSYRCKCRDGNB 720
 DB 684 SLIHVAAYADLRHYHODVLEMLCGEAKOPVNLCKSPCMNBGSCTVLQNGSYRCKCRDGNB 743
 QY 721 GPHCENRFLRRP 732
 DB 744 GPHCENRFLRRP 755

RESULT 14
 US-09-930-020A-2
 / Sequence 2, Application US/09930020A
 / Publication No. US20030077568A1
 / GENERAL INFORMATION:
 / APPLICANT: Gish, Kurt C.
 / APPLICANT: Mack, David H.
 / APPLICANT: Wilson, Keith E.
 / APPLICANT: Eos Biotechnology, Inc.
 / TITLE OF INVENTION: Methods of diagnosis of Colorectal Cancer, Compositions
 / TITLE OF INVENTION: and Methods of Screening for Colorectal Cancer
 / TITLE OF INVENTION: Modulators
 / FILE REFERENCE: 018501-003100US
 / CURRENT APPLICATION NUMBER: US/09/930,020A
 / PRIOR FILING DATE: 2001-08-14
 / PRIOR APPLICATION NUMBER: US 09/663,733
 / PRIOR FILING DATE: 2000-09-15
 / NUMBER OF SEQ ID NOS: 3
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 2
 / LENGTH: 807
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 / FEATURE:
 / OTHER INFORMATION: CBF9
 US-09-930-020A-2

Query Match 99.3%; Score 3799; DB 3; Length 807;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHNSKETIGKISAKMMKCSAAVDIMFLDGSNSVKGSPERSKHPAIVTCGGLD 60
 DB 24 LOEVHNSKETIGKISAKMMKCSAAVDIMFLDGSNSVKGSPERSKHPAIVTCGGLD 83
 QY 61 SPERRVGAFOGSSPTLPEPLDLSFSTQOEYKARIKRMVKGKRTTELAKTLILRGUP 120
 DB 84 SPERRVGAFOGSSPTLPEPLDLSFSTQOEYKARIKRMVKGKRTTELAKTLILRGUP 143
 QY 121 GGNNAVPOQLIVTDOGKSGDVALPSKQKRGVTVFAVGVPFPMWELHALASEPRQ 180

DB 144 GGRNAVPOQLIVTDOGKSGDVALPSKQKRGVTVFAVGVPFPMWELHALASEPRQ 203
 QY 181 HYLLAEQVEDATNGFLSTSSAICSSATPDQCRVAHPPEBHNTLEWRFPAQNAPEWRS 240
 DB 204 HYLLAEQVEDATNGFLSTSSAICSSATPDQCRVAHPPEBHNTLEWRFPAQNAPEWRS 263
 QY 241 RRTLVLAHCPFYSGKRVFLTHPATCYRTTCGPPCDSDPCCONGGTCVEGDLGYOCLCP 300
 DB 264 RRTLVLAHCPFYSGKRVFLTHPATCYRTTCGPPCDSDPCCONGGTCVEGDLGYOCLCP 323
 QY 301 LAFGEANCALKLSLECRVDLFLDSSAGTTLDGFLRAKVFYKRFVRAVLSEDSBARVG 360
 DB 324 LAFGEANCALKLSLECRVDLFLDSSAGTTLDGFLRAKVFYKRFVRAVLSEDSBARVG 383
 QY 361 VATTYRELLVAAPVGEYQDPDLVMSLDGI PRGPGPTLTGSALRQAERGFSGATRTGDD 420
 DB 384 VATTYRELLVAAPVGEYQDPDLVMSLDGI PRGPGPTLTGSALRQAERGFSGATRTGDD 443
 QY 421 RRRRVVLLTSHSEDEBVAQPARHARARELLLGVSSEAVRALEBETITGSPKRWVYSDP 480
 DB 444 RRRRVVLLTSHSEDEBVAQPARHARARELLLGVSSEAVRALEBETITGSPKRWVYSDP 503
 QY 481 ODLENOIPELOGLCSRGPRGCRTOALDVFMLDTSASVGPENFAQMOGFSVSCALQFEV 540
 DB 504 ODLENOIPELOGLCSRGPRGCRTOALDVFMLDTSASVGPENFAQMOGFSVSCALQFEV 563
 QY 541 NPDVTQGLVYVGSQVQTAFGDTPKTRAMLRASQAAPYLGVSAGTALLHIYDKWNT 600
 DB 564 NPDVTQGLVYVGSQVQTAFGDTPKTRAMLRASQAAPYLGVSAGTALLHIYDKWNT 623
 QY 601 VORGARPGVPKAVVYLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 660
 DB 624 VORGARPGVPKAVVYLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSBGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHYHODVLEMLCGEAKOPVNLCKSPCMNBGSCTVLQNGSYRCKCRDGNB 720
 DB 684 SLIHVAAYADLRHYHODVLEMLCGEAKOPVNLCKSPCMNBGSCTVLQNGSYRCKCRDGNB 743
 QY 721 GPHCENR 727
 DB 744 GPHCENR 750

RESULT 15
 US-10-295-027-122
 / Sequence 122, Application US/10295027
 / Publication No. US20030232350A1
 / GENERAL INFORMATION:
 / APPLICANT: Afari, Daniel
 / APPLICANT: Aziz, Natsaana
 / APPLICANT: Ginsberg, Wendy M.
 / APPLICANT: Gish, Kurt C.
 / APPLICANT: Glynn, Richard
 / APPLICANT: Hevezl, Peter A.
 / APPLICANT: Mack, David H.
 / APPLICANT: Murray, Richard
 / APPLICANT: Watson, Susan R.
 / APPLICANT: Eos Biotechnology, Inc.
 / TITLE OF INVENTION: Methods of diagnosis of Cancer, Compositions and
 / TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
 / FILE REFERENCE: 018501-012500US
 / CURRENT APPLICATION NUMBER: US/10/295,027
 / PRIOR FILING DATE: 2002-11-13
 / PRIOR APPLICATION NUMBER: US 09/663,733
 / PRIOR FILING DATE: 2000-09-15
 / PRIOR APPLICATION NUMBER: US 60/350,666
 / PRIOR FILING DATE: 2001-11-13
 / PRIOR APPLICATION NUMBER: US 60/335,394
 / PRIOR FILING DATE: 2001-11-15
 / PRIOR APPLICATION NUMBER: US 60/332,464
 / PRIOR FILING DATE: 2001-11-21
 / PRIOR APPLICATION NUMBER: US 60/334,393

1 PRIOR FILING DATE: 2001-11-29
1 PRIOR APPLICATION NUMBER: US 60/340,376
1 PRIOR FILING DATE: 2001-12-14
1 PRIOR APPLICATION NUMBER: US 60/347,211
1 PRIOR FILING DATE: 2002-01-08
1 PRIOR APPLICATION NUMBER: US 60/347,349
1 PRIOR FILING DATE: 2002-01-10
1 PRIOR APPLICATION NUMBER: US 60/355,250
1 PRIOR FILING DATE: 2002-02-08
1 PRIOR APPLICATION NUMBER: US 60/356,714
1 PRIOR FILING DATE: 2002-02-13
1 Remaining Prior Application data removed - See File Wrapper or PALM.
1 NUMBER OF SEQ ID NOS: 1386
1 SOFTWARE: PatentIn Ver. 2.1
1 SEQ ID NO 122
1 LENGTH: 807
1 TYPE: PRT
1 ORGANISM: Homo sapiens
US-10-295-027-122

Query Match 99.3%; Score 3799; DB 4; Length 807;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKETIGKISAAKMMCSAAVDIMFLDGSNSVGKGFERSKHPAIVCDGLDI 60
DB 24 LOEVHVSKETIGKISAAKMMCSAAVDIMFLDGSNSVGKGFERSKHPAIVCDGLDI 83
QY 61 SPERVAVGAFQFSTSTPHLEFPDLSFSTQDEVKARIKRMVFKGRTETELAKYLHRLP 120
DB 84 SPERVAVGAFQFSTSTPHLEFPDLSFSTQDEVKARIKRMVFKGRTETELAKYLHRLP 143
QY 121 GGNASVPOLLIIIVTDKSGQDVALLPSKQIKERGVTFAVGVPFPMEEIHALASEPRQ 180
DB 144 GGNASVPOLLIIIVTDKSGQDVALLPSKQIKERGVTFAVGVPFPMEEIHALASEPRQ 203
QY 181 HVLAAQVEDATNGLFSTLSSSAICSSATPDCEVAHPCEHRTLEMRBEPAGNAPCMRGS 240
DB 204 HVLAAQVEDATNGLFSTLSSSAICSSATPDCEVAHPCEHRTLEMRBEPAGNAPCMRGS 263
QY 241 RRTLAIVLAHCPFYSMKRVFLTHPATCYRTTCGPGCSOPCONGTCPVEGLDGYQCLCP 300
DB 264 RRTLAIVLAHCPFYSMKRVFLTHPATCYRTTCGPGCSOPCONGTCPVEGLDGYQCLCP 323
QY 301 LAFGSEANCAKLISLDCRVLDLFLDSSAGTTLDFLRAKVPYKRFVRAVLSSEDSRAVG 360
DB 324 LAFGSEANCAKLISLDCRVLDLFLDSSAGTTLDFLRAKVPYKRFVRAVLSSEDSRAVG 383
QY 361 VATYSRELVAVPVGVGYQVDPDLVMSLDGIPFRGSPITLGSALROAERFGSATRTGDD 420
DB 384 VATYSRELVAVPVGVGYQVDPDLVMSLDGIPFRGSPITLGSALROAERFGSATRTGDD 443
QY 421 RPRRVVLLTSHSEDEVAAPAHARARELLILGVSEAVRALEBITGSPKHVMYSDP 480
DB 444 RPRRVVLLTSHSEDEVAAPAHARARELLILGVSEAVRALEBITGSPKHVMYSDP 503
QY 481 ODLENGIPELQKLCGRQRPGRCTQALDLVFMLDTSASVGPENFAQMGFVRSICALQFEV 540
DB 504 ODLENGIPELQKLCGRQRPGRCTQALDLVFMLDTSASVGPENFAQMGFVRSICALQFEV 563
QY 541 NPDVNOVGLVVGSGOYQTAFGDLTKPTRAAMLRAISOAPYLGVGSAGTALLHIYDKVMT 600
DB 564 NPDVNOVGLVVGSGOYQTAFGDLTKPTRAAMLRAISOAPYLGVGSAGTALLHIYDKVMT 623
QY 601 VORGARPGVPAKVVVLTGGRGAEDAAPPAOKLRNNGISVTVGVGPVLSSEGLRRLAGPRD 660
DB 624 VORGARPGVPAKVVVLTGGRGAEDAAPPAOKLRNNGISVTVGVGPVLSSEGLRRLAGPRD 683
QY 661 SLIHVAAYADLRYHQDVLLEMLGSAKQPVNLCKPSPCMNNGSCVYLQNGSYRCKCRDGE 720
DB 684 SLIHVAAYADLRYHQDVLLEMLGSAKQPVNLCKPSPCMNNGSCVYLQNGSYRCKCRDGE 743
QY 721 GPHCENR 727

Db 744 GPHCENR 750
Search completed: June 21, 2006, 20:28:48
Job time : 183 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2006, 20:04:13 ; Search time 201 Seconds

(without alignments)
1665.087 Million cell updates/sec

Title: US-10-649-591-3

3826

Sequence: 1 LQEVHVSKEITGKISASKM.....CKCRDGMGPHCNFLARP 732

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_8:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*
10: geneseqp2006s:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3826	100.0	732	8	ADL70216
2	3826	100.0	732	8	ADN59608
3	3826	100.0	755	6	ABUS6623
4	3826	100.0	755	7	ADH80523
5	3826	100.0	755	7	ADN38806
6	3826	100.0	755	8	ADL70227
7	3826	100.0	755	8	ADN59619
8	3826	99.9	755	7	ADN39868
9	3799	99.3	807	5	ABG61891
10	3799	99.3	807	5	ABP54687
11	3799	99.3	807	6	ABUS6622
12	3799	99.3	807	7	ADH80521
13	3799	99.3	807	7	ADN39869
14	3799	99.3	807	7	ADN38804
15	3799	99.3	807	7	ADN39446
16	3799	99.3	807	7	ADN39459
17	3799	99.3	807	7	ADN39565
18	3799	99.3	807	8	ADO70146
19	3540	92.5	725	8	ADQ65732
20	3528	92.2	725	8	ADR09513
21	3387.5	88.5	1160	4	ABG17343
22	3382.5	88.4	826	6	ABUS6621
23	3382.5	88.4	826	7	ADH80519

24	3382.5	88.4	826	7	ADN38802	Adn38802	Cancer/an
25	3177.5	83.1	731	8	ADG78854	Adg78854	Human can
26	2568	67.1	488	8	ADL70234	Adl70234	Colon can
27	2568	67.1	488	8	ADN59626	Adn59626	Colon neo
28	1580	41.3	405	3	AAV76095	Aay76095	Murine bk
29	1580	41.3	405	3	AAAB56034	Aab56034	Skin cell
30	1580	41.3	405	5	ABB72234	Abb72234	Murine pr
31	522.5	13.7	2000	9	ADP54237	Adp54237	Human PRO
32	522.5	13.7	2000	9	ADY14520	Ady14520	PRO polyP
33	522.5	13.7	2211	8	ADQ39661	Adq39661	Human myo
34	522.5	13.7	2244	8	ADQ39658	Adq39658	Human myo
35	522.5	13.7	2411	8	ADQ39660	Adq39660	Human myo
36	522.5	13.7	2977	8	ABW82566	Abw82566	Human dia
37	522.5	13.7	2977	8	ADQ39657	Adq39657	Human myo
38	522.5	13.7	3151	6	ABR39571	Ab39571	Human alp
39	522.5	13.7	3173	5	ABB90743	Abb90743	Human tum
40	522.5	13.7	3173	6	ABU54450	Abu54450	Human tum
41	522.5	13.7	3176	6	ABR39570	Ab39570	Human alp
42	522.5	13.7	3176	7	ADP65255	Adp65255	Human alp
43	522.5	13.7	3176	8	ADQ18247	Adq18247	Human sof
44	522.5	13.7	3176	8	ADQ29665	Adq29665	Human col
45	522.5	13.7	3176	8	ADR87616	Adr87616	Human typ

ALIGNMENTS

RESULT 1
ID ADL70216 standard; protein; 732 AA.
XX
AC ADL70216;
XX
DT 20-MAY-2004 (first entry)
XX
DE Colon cancer marker ColoUp2 secreted protein.
XX
KM ColoUp2; colon cancer; marker; diagnosis; cytostatic; vaccine;
XX gene therapy.
XX
OS Homo sapiens.
XX
PN WO2004018648-A2.
XX
PD 04-MAR-2004.
XX
PF 26-AUG-2003; 2003WO-US027086.
XX
PR 26-AUG-2002; 2002US-00229345.
XX
PR 27-AUG-2002; 2002US-0406296P.
XX
PR 18-OCT-2002; 2002US-002741177.
XX
PA (UYCA-) UNIV CASE WESTERN RESERVE.
XX
PI Markowitz SD;
XX
DR WPI; 2004-226834/21.
XX
DR N-PSDB; ADL70218.
XX
PT Inhibiting the growth or proliferation of a colon neoplasia in a subject,
XX for treating colon cancer, comprises administering to the subject an
XX agent that decreases the amount of a polypeptide present in or produced
XX by the colon neoplasia.
XX
PS Claim 10; SEQ ID NO 3; 125pp; English.
XX
CC The present sequence is the protein sequence of secreted ColoUp2, a
XX molecular marker of colon neoplasia. ColoUp2 was identified by expression
XX microarray profiling of genes showing differential expression in normal
XX colon and metastatic colon cancer. ColoUp2 transcript is essentially
XX undetectable in normal colon epithelial strips, in normal and in colonic
XX muscle, but expression is clearly detectable in premalignant colon
XX adenomas and in 90% of Duke's stage B (early negative colon cancers),

CC Dukes stage C (node positive colon cancer), Dukes stage D (primary colon
 CC cancers with associated metastatic spread) and in colon cancer liver
 CC metastasis. Expression is also detectable in colon cancer cell lines and
 CC in colon cancer xenografts grown in athymic mice. Col02p2 is secreted
 CC from both the basolateral and apical surfaces of intestinal cells, and
 CC can be found in the blood in 2 different forms, a full-length secreted
 CC form and a C-terminal fragment. The molecular marker is useful in the
 CC detection or diagnosis of colon neoplasia and for categorising the
 CC neoplastic state of a patient. Methods for inhibiting the growth or
 CC proliferation of a colon neoplasia involve the use of host interfering
 CC RNA (siRNA) or antisense oligonucleotides, or an agent that binds to and
 CC antagonises the polypeptide. Molecular markers can also be used to target
 CC chemotherapeutic agents to cells of a colon neoplasia, and to screen and
 CC identify candidate therapeutic agents.

XX
 XX
 SQ Sequence 732 AA;

Query Match 100.0%; Score 3826; DB 8; Length 732;
 Best Local Similarity 100.0%; Pred. No. 6, 3e-300;
 Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEVHVSKEITIGKISAAKMMCSAAVDIMFLDGSNSVGKSFERSKGFATVCDGLDI 60
 DB 1 LQEVHVSKEITIGKISAAKMMCSAAVDIMFLDGSNSVGKSFERSKGFATVCDGLDI 60
 QY 61 SPERVVGAFOFSSSTPHLEFPDLSFTQOEYKARIKRWVFKGRTETELAKYLHARGLP 120
 DB 61 SPERVVGAFOFSSSTPHLEFPDLSFTQOEYKARIKRWVFKGRTETELAKYLHARGLP 120
 QY 121 GGRNASVPOILITVTDGKSQGVVALPSKQLKERGVTVFAVGVPFRMBELHALASEPRQ 180
 DB 121 GGRNASVPOILITVTDGKSQGVVALPSKQLKERGVTVFAVGVPFRMBELHALASEPRQ 180
 QY 181 HYLAEQVEDATNGLFSTLSSSAICSSATPDCEVAHPCERHTELVAREFAGNAPCWRGS 240
 DB 181 HYLAEQVEDATNGLFSTLSSSAICSSATPDCEVAHPCERHTELVAREFAGNAPCWRGS 240
 QY 241 RRTLAVALAHCPEYSMKRVFLTHPATCYRTTCGPGCDSPQCONGTCTVEBGLDGYQCLCP 300
 DB 241 RRTLAVALAHCPEYSMKRVFLTHPATCYRTTCGPGCDSPQCONGTCTVEBGLDGYQCLCP 300
 QY 301 LAFGGANCAKLKLSLECRVDLFLDSSAKGTTLDGFLRAKVPYKRVRAVLSEDSRAVVG 360
 DB 301 LAFGGANCAKLKLSLECRVDLFLDSSAKGTTLDGFLRAKVPYKRVRAVLSEDSRAVVG 360
 QY 361 VATYSRELLVAVPVGEYQVDPDLVMSLDGIPFRGSGTTLGSAALRQAAERGFSGSATRTGQD 420
 DB 361 VATYSRELLVAVPVGEYQVDPDLVMSLDGIPFRGSGTTLGSAALRQAAERGFSGSATRTGQD 420
 QY 421 RPRRVVVLLTSHSEDEVAAPARHARARELLLGVSSEAVRALEBITGSPKRWVYSDP 480
 DB 421 RPRRVVVLLTSHSEDEVAAPARHARARELLLGVSSEAVRALEBITGSPKRWVYSDP 480
 QY 481 QDLFNOIPELQKLCGRQRPGRCTQALDIVPMLDTSASGPNFPMQSFVSSCALQFVY 540
 DB 481 QDLFNOIPELQKLCGRQRPGRCTQALDIVPMLDTSASGPNFPMQSFVSSCALQFVY 540
 QY 541 NPDVTOGVAVVYGSQVQTAFGDLTKRPAAMLRASQAPYLGGVGSAGTALLHITDKWMT 600
 DB 541 NPDVTOGVAVVYGSQVQTAFGDLTKRPAAMLRASQAPYLGGVGSAGTALLHITDKWMT 600
 QY 601 VORGARPGVPAVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSSEGLRRLAGPRD 660
 DB 601 VORGARPGVPAVVVLTGGRGAEDAAVPAQKLRNNGISVLYVGVGVLSSEGLRRLAGPRD 660
 QY 661 SLIHYAAVADLRHNOVLLEMLGGEAKQPVNLCKPSPCMNEGSCVLYQNSYCKCRDGE 720
 DB 661 SLIHYAAVADLRHNOVLLEMLGGEAKQPVNLCKPSPCMNEGSCVLYQNSYCKCRDGE 720
 QY 721 GPHCENRFLRRP 732
 DB 721 GPHCENRFLRRP 732

RESULT 2
 ADNS9608
 ID ADNS9608 brandard; protein; 732 AA.

XX AC ADNS9608;

XX DT 01-JUL-2004 (first entry)

XX DE Colon neoplasia Col02p2 protein, SEQ ID No 3.

XX KW molecular marker; neoplastic state; colon neoplasia; Col02p1; Col02p2;
 KW colon cancer.

XX OS Unidentified.

XX PN WO2004018647-A2.

XX PD 04-MAR-2004.

XX PF 26-AUG-2003; 2003WO-US027085.

XX PR 26-AUG-2002; 2002US-00229345.

XX PR 18-OCT-2002; 2002US-00274177.

XX PA (UYCA-) UNIV CASE WESTERN RESERVE.

XX PI Markowitz SD;

XX DR WPI, 2004-419536/39.

XX PT Detecting likelihood of subject to have colon neoplasia, involves
 PT obtaining biological sample from subject, detecting one or more of
 PT Col02p1 and Col02p2 polypeptides, presence of polypeptide indicating
 PT colon neoplasia.

XX PS Claim 11; SEQ ID NO 3; 118bp; English.

XX SQ Sequence 732 AA;

CC The invention relates to novel molecular markers for categorizing the
 CC neoplastic state of a patient. The invention further comprises a method
 CC for detecting whether a subject is likely to have a colon neoplasia. The
 CC method comprises: obtaining a biological sample from the subject;
 CC detecting one or more polypeptides selected from among: one or more
 CC Col02p1 polypeptides and one or more Col02p2 polypeptides, wherein the
 CC presence of said one or more polypeptides is indicative of colon
 CC neoplasia. The invention further comprises a kit for carrying out the
 CC said method. The methods and compositions are useful for detecting colon
 CC cancer. This sequence represents a Col02p2 protein of the invention.

Query Match 100.0%; Score 3826; DB 8; Length 732;
 Best Local Similarity 100.0%; Pred. No. 6, 3e-300;
 Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEVHVSKEITIGKISAAKMMCSAAVDIMFLDGSNSVGKSFERSKGFATVCDGLDI 60
 DB 1 LQEVHVSKEITIGKISAAKMMCSAAVDIMFLDGSNSVGKSFERSKGFATVCDGLDI 60
 QY 61 SPERVVGAFOFSSSTPHLEFPDLSFTQOEYKARIKRWVFKGRTETELAKYLHARGLP 120
 DB 61 SPERVVGAFOFSSSTPHLEFPDLSFTQOEYKARIKRWVFKGRTETELAKYLHARGLP 120
 QY 121 GGRNASVPOILITVTDGKSQGVVALPSKQLKERGVTVFAVGVPFRMBELHALASEPRQ 180
 DB 121 GGRNASVPOILITVTDGKSQGVVALPSKQLKERGVTVFAVGVPFRMBELHALASEPRQ 180
 QY 181 HYLAEQVEDATNGLFSTLSSSAICSSATPDCEVAHPCERHTELVAREFAGNAPCWRGS 240
 DB 181 HYLAEQVEDATNGLFSTLSSSAICSSATPDCEVAHPCERHTELVAREFAGNAPCWRGS 240
 QY 241 RRTLAVALAHCPEYSMKRVFLTHPATCYRTTCGPGCDSPQCONGTCTVEBGLDGYQCLCP 300
 DB 241 RRTLAVALAHCPEYSMKRVFLTHPATCYRTTCGPGCDSPQCONGTCTVEBGLDGYQCLCP 300

QY 601 VORGARPGVPKAVVLTGGGAEADAAPQAQKLNNNGISVLTVGVGVLTSEGLRRLAGPRD 660
 DB 624 VORGARPGVPKAVVLTGGGAEADAAPQAQKLNNNGISVLTVGVGVLTSEGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHYHDLIEMTCGEAKOPVNLCKPSPCMNBSGCVLQNGSYRCKCRDGM 720
 DB 684 SLIHVAAYADLRHYHDLIEMTCGEAKOPVNLCKPSPCMNBSGCVLQNGSYRCKCRDGM 743
 QY 721 GPHCENRFLRRP 732
 DB 744 GPHCENRFLRRP 755

RESULT 4
 ADB80523
 ID ADB80523 standard; protein; 755 AA.
 XX ADB80523;
 AC ADB80523;
 DT 04-DEC-2003 (first entry)
 XX
 DE Ovarian cancer-associated protein #44.
 XX
 KW cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
 KW post-operative chemotherapy; radiation therapy; tumour prognosis;
 KW pre-cancerous lesion detection.
 XX
 OS Homo sapiens.
 XX
 PN MO2002102235-AA.
 PD 27-DEC-2002.
 XX
 PF 18-JUN-2002; 2002MO-US019297.
 XX
 PR 18-JUN-2001; 2001US-0299234P.
 PR 27-AUG-2001; 2001US-0315287P.
 PR 05-SEP-2001; 2001US-0317544P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 PI Mack DH, Glash KC;
 XX
 DR MPI; 2003-167431/16.
 N-PSDB; ADB80522.
 PT Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.
 XX
 PS Claim 13; Page 304-305; 332pp; English.
 XX
 CC The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the proteins used
 CC for the detection method of the invention.
 CC
 XX
 SQ Sequence 755 AA;
 Query Match 100.0%; Score 3826; DB 7; Length 755;
 Best Local Similarity 100.0%; Pred. No. 6.6e-300;

Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LQEVHVSKEETIKISAASPMWCSAAVDIMFLDGSNSVKGKSPERSKFAITVCGLDI 60
 DB 24 LQEVHVSKEETIKISAASPMWCSAAVDIMFLDGSNSVKGKSPERSKFAITVCGLDI 83
 QY 61 SPERRVGAFOFSSTPHLEFPIDSFSTQOEVKARIKRMVFKGRTETELAKYLLHRGLP 120
 DB 84 SPERRVGAFOFSSTPHLEFPIDSFSTQOEVKARIKRMVFKGRTETELAKYLLHRGLP 143
 QY 121 GGRNASVPQIILITVDGKSQGDVALPSKQIKRGVTVPVGVFRPMEHLHALASPRQ 180
 DB 144 GGRNASVPQIILITVDGKSQGDVALPSKQIKRGVTVPVGVFRPMEHLHALASPRQ 203
 QY 181 HTVLAEOVEDATNGLPSTLSSAICSSATPDGRVAHPCEHNTLEMYREFPAGNAPCMRGS 240
 DB 204 HTVLAEOVEDATNGLPSTLSSAICSSATPDGRVAHPCEHNTLEMYREFPAGNAPCMRGS 263
 QY 241 RRTTLAVLAHCPFYSMKRVFLTHPATCYRTTCPRGCDSPQCONGGTCVBEGLDGYOCLCP 300
 DB 264 RRTTLAVLAHCPFYSMKRVFLTHPATCYRTTCPRGCDSPQCONGGTCVBEGLDGYOCLCP 323
 QY 301 IAFGEANCAKLSLECGVDLITLDSAGTLDGFLAKVVKRFPVAVVSEDSARVG 360
 DB 324 IAFGEANCAKLSLECGVDLITLDSAGTLDGFLAKVVKRFPVAVVSEDSARVG 383
 QY 361 VATYREBLVAVPVEGYDVPDLVWSLDGIPFRGGPTLTGSLARQAAERGFSGATRTGQD 420
 DB 384 VATYREBLVAVPVEGYDVPDLVWSLDGIPFRGGPTLTGSLARQAAERGFSGATRTGQD 443
 QY 421 RPRRVVLTITSHSDEVAAGPARHARARELLLVGSEAVNALEBITSPKGVWVYSDP 480
 DB 444 RPRRVVLTITSHSDEVAAGPARHARARELLLVGSEAVNALEBITSPKGVWVYSDP 503
 QY 481 QDLFNOIPELOKLSRORPGCRTLADLVFMLDTSASVGPENFPMOMSFVNSCALQEPV 540
 DB 504 QDLFNOIPELOKLSRORPGCRTLADLVFMLDTSASVGPENFPMOMSFVNSCALQEPV 563
 QY 541 NPDVTVQVGLVYVGSQVQTAFGIDTKPTRAAMLRAISQAPYLGVSAGTALHIDYKWT 600
 DB 564 NPDVTVQVGLVYVGSQVQTAFGIDTKPTRAAMLRAISQAPYLGVSAGTALHIDYKWT 623
 QY 601 VORGARPGVPKAVVLTGGGAEADAAPQAQKLNNNGISVLTVGVGVLTSEGLRRLAGPRD 660
 DB 624 VORGARPGVPKAVVLTGGGAEADAAPQAQKLNNNGISVLTVGVGVLTSEGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHYHDLIEMTCGEAKOPVNLCKPSPCMNBSGCVLQNGSYRCKCRDGM 720
 DB 684 SLIHVAAYADLRHYHDLIEMTCGEAKOPVNLCKPSPCMNBSGCVLQNGSYRCKCRDGM 743
 QY 721 GPHCENRFLRRP 732
 DB 744 GPHCENRFLRRP 755

RESULT 5
 ADN38806
 ID ADN38806 standard; protein; 755 AA.
 XX ADN38806;
 AC ADN38806;
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:124.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; porriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KW vulnerary; gene therapy; vaccine.
 XX

OS Homo sapiens.
 XX WO2003042661-A2.
 XX
 XX
 XX
 PD 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-038614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397755P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA
 PI Afar D, Aziz N, Gineburg WM, Gish KC, Glynn R, Hevezl PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;
 XX
 XX WPI; 2003-468649/44.
 DR N-PSDB; ADN38805.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 XX a nucleic acid in a biological sample.
 XX
 PS Claim 12; SEQ ID NO 124; 1385bp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.
 XX
 XX Sequence 755 AA;

Query Match 100.0%; Score 3826; DB 7; Length 755;
 Best Local Similarity 100.0%; Pred. No. 6,6e-300; Indels 0; Gaps 0;
 Matches 732; Conservative 0; Mismatches 0;

QY 1 LOEVHYSKETIGTISAASKMMSCAAVDIMFLDGSNSVGSGPERSKHPAIVTCGLDI 60
 DB 24 LOEVHYSKETIGTISAASKMMSCAAVDIMFLDGSNSVGSGPERSKHPAIVTCGLDI 83
 QY 61 SPERVAVGAFQFSSTPHLEFPDLSFSTQEVKARIKMWFKGRTETELAKTLHRLGP 120
 DB 84 SPERVAVGAFQFSSTPHLEFPDLSFSTQEVKARIKMWFKGRTETELAKTLHRLGP 143

QY 121 GGRNASVPOIILIIYTDGKSQGDVALPSKQLKRGVTVFAVGRPRMELHALASEPRGQ 180
 DB 144 GGRNASVPOIILIIYTDGKSQGDVALPSKQLKRGVTVFAVGRPRMELHALASEPRGQ 203
 QY 181 HTVLAEQVEDATNGLPSTLISSAICSSATPDCCREVAHPCEHRTLEKMEFAFNAPCMRGS 240
 DB 204 HTVLAEQVEDATNGLPSTLISSAICSSATPDCCREVAHPCEHRTLEKMEFAFNAPCMRGS 263
 QY 241 RRTLVLAHCPFPYSMKVFLTHPATCYRTTCRPGCDSPQONGGTCVPEGIDGYQCLCP 300
 DB 264 RRTLVLAHCPFPYSMKVFLTHPATCYRTTCRPGCDSPQONGGTCVPEGIDGYQCLCP 323
 QY 301 IAFGEANCAIKLISECRVDLFLDSSAGTTLDFLPAKVVKKEFVAVASEDSRARVG 360
 DB 324 IAFGEANCAIKLISECRVDLFLDSSAGTTLDFLPAKVVKKEFVAVASEDSRARVG 383
 QY 361 VATTYRELLVAVPVEEYQDPDLVMSLDGIPPRGPTLTGSLAKQDAERGFSSATRTGQD 420
 DB 384 VATTYRELLVAVPVEEYQDPDLVMSLDGIPPRGPTLTGSLAKQDAERGFSSATRTGQD 443
 QY 421 RRRRVVLLTSSHSDDEVAGPARHARRELLILGVSEAVRALEBEITGSPKHWVYSDP 480
 DB 444 RRRRVVLLTSSHSDDEVAGPARHARRELLILGVSEAVRALEBEITGSPKHWVYSDP 503
 QY 481 QDLFNOIPELQKLCSPRPGCRTOALDLVFMLDTSASVGPENFAQOMOSFVNSCALQFEV 540
 DB 504 QDLFNOIPELQKLCSPRPGCRTOALDLVFMLDTSASVGPENFAQOMOSFVNSCALQFEV 563
 QY 541 NPDVTQVGLVYVGSQVQTAFLGDTKPTRAMLRASQADYLGCVSAGTALLHIYDKVMT 600
 DB 564 NPDVTQVGLVYVGSQVQTAFLGDTKPTRAMLRASQADYLGCVSAGTALLHIYDKVMT 623
 QY 601 VORGARPGVPAVWVLTGCRGAEDAAVPAQKLRNNGISVLYVGVDPVYSEGLRRLAGPRD 660
 DB 624 VORGARPGVPAVWVLTGCRGAEDAAVPAQKLRNNGISVLYVGVDPVYSEGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHYDVLIEMLGCAKQPVNLCKPSPCMNBSGCVLQNSYRCKCRDGE 720
 DB 684 SLIHVAAYADLRHYDVLIEMLGCAKQPVNLCKPSPCMNBSGCVLQNSYRCKCRDGE 743
 QY 721 GPHCENRFLRRP 732
 DB 744 GPHCENRFLRRP 755

RESULT 6
 ADL70227
 ID ADL70227 standard; protein; 755 AA.
 XX
 AC ADL70227;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Colon cancer marker Colo2p2 full-length protein.
 XX
 KW Colo2p2; colon cancer; marker; diagnosis; cytostatic; vaccine;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein 24..732
 FT /label= "Signal peptide
 FT /label= "Secreted protein
 FT /note= "Also referred to in Claim 10"
 XX
 XX MO2004018648-A2.
 XX
 PD 04-MAR-2004.
 XX
 PF 26-AUG-2003; 2003WO-US027086.
 XX

PR 26-AUG-2002; 2002US-00229345.
 PR 27-AUG-2002; 2002US-0406296P.
 PR 18-OCT-2002; 2002US-00274177.
 XX
 XX
 PA (UYCA-) UNIV CASE WESTERN RESERVE.
 XX
 PI Markowitz SD;
 XX
 DR WPI; 2004-226834/21.
 XX
 PT Inhibiting the growth or proliferation of a colon neoplasia in a subject,
 PT for treating colon cancer, comprises administering to the subject an
 PT agent that decreases the amount of a polypeptide present in or produced
 PT by the colon neoplasia.
 XX
 PS Claim 10; SEQ ID NO 14; 125bp; English.
 XX
 CC The present sequence is the protein sequence of full-length ColoUp2, a
 CC molecular marker of colon neoplasia. ColoUp2 was identified by expression
 CC microarray profiling of genes showing differential expression in normal
 CC colon and metastatic colon cancer. ColoUp2 transcript is essentially
 CC undetectable in normal colon epithelial strips, in normal and in colonic
 CC muscle, but expression is clearly detectable in premalignant colon
 CC adenomas and in 90% of Duke's stage B (early negative colon cancers),
 CC Duke's stage C (node positive colon cancer), Duke's stage D (primary colon
 CC cancers with associated metastatic spread) and in colon cancer liver
 CC metastasis. Expression is also detectable in colon cancer cell lines and
 CC in colon cancer xenografts grown in athymic mice. ColoUp2 is secreted
 CC from both the basolateral and apical surfaces of intestinal cells, and
 CC can be found in the blood in 2 different forms, a full-length secreted
 CC form and a C-terminal fragment. The molecular marker is useful in the
 CC detection or diagnosis of colon neoplasia and for categorizing the
 CC neoplastic state of a patient. Methods for inhibiting the growth or
 CC proliferation of a colon neoplasia involve the use of short interfering
 CC RNA (siRNA) or antisense oligonucleotides, or an agent that binds to and
 CC antagonizes the polypeptide. Molecular markers can also be used to target
 CC therapeutic agents to cells of a colon neoplasia, and to screen and
 CC identify candidate therapeutic agents.
 XX
 XX Sequence 755 AA;
 SQ
 Query Match 100.0%; Score 3826; DB 8; Length 755;
 Best Local Similarity 100.0%; Pred. No. 6, 6e-300;
 Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LQEVHVSKEITGKISNASKMMCSAAVDIMFLDGSNSVKGSGFERSKHFATVCGDLDI 60
 DB 24 LQEVHVSKEITGKISNASKMMCSAAVDIMFLDGSNSVKGSGFERSKHFATVCGDLDI 83
 QY 61 SPBRVAVGAFQSSSTPHLEFPDLSFSTQGVKARIKRMVFKGRTETELAKYLLHRLGP 120
 DB 84 SPBRVAVGAFQSSSTPHLEFPDLSFSTQGVKARIKRMVFKGRTETELAKYLLHRLGP 143
 QY 121 GGRNASVPOILIIYTDGKSGQGVALPSKQLKERGVAVFVGVFRFPWEELHALASPRQ 180
 DB 144 GGRNASVPOILIIYTDGKSGQGVALPSKQLKERGVAVFVGVFRFPWEELHALASPRQ 203
 QY 181 HYLAAQVEDATNGLFSTLSSAICSSATPDCRVEAHPCERHTELEWREAFAGNAPCRMS 240
 DB 204 HYLAAQVEDATNGLFSTLSSAICSSATPDCRVEAHPCERHTELEWREAFAGNAPCRMS 263
 QY 241 RRTLAVLAHCPFYSKRVFLTHPATCYTTTCGPGCDSPPCONGTCVPEGLDGYOCLCP 300
 DB 264 RRTLAVLAHCPFYSKRVFLTHPATCYTTTCGPGCDSPPCONGTCVPEGLDGYOCLCP 323
 QY 301 LAFGGANCAKLKLSIECRVDLLFLDSSAGTLLDGLFAKVVFYKRVRAVLSEDSPARVG 360
 DB 324 LAFGGANCAKLKLSIECRVDLLFLDSSAGTLLDGLFAKVVFYKRVRAVLSEDSPARVG 383
 QY 361 VATYSRELLVAVPVGEYQVDPDLVMSLDGI PRFGGPTLLTGSALRQAAERGFSSATRTGDD 420
 DB 384 VATYSRELLVAVPVGEYQVDPDLVMSLDGI PRFGGPTLLTGSALRQAAERGFSSATRTGDD 443

QY 421 RPRRVVVLITESHSEDEVAGPARHARARELLILGVGSEAVRAELEBITGSPRHVMYSDP 480
 DB 444 RPRRVVVLITESHSEDEVAGPARHARARELLILGVGSEAVRAELEBITGSPRHVMYSDP 503
 QY 481 QDLFNOIPELOGLKLSRORPGCRTOALDVFLMDTSASVGPENFOMOSFVNSCALQFHV 540
 DB 504 QDLFNOIPELOGLKLSRORPGCRTOALDVFLMDTSASVGPENFOMOSFVNSCALQFHV 563
 QY 541 NPDVTQVGLVYVGSQVOTAFGLDTKPTRAAMLRAISOAPVYLGVGSAGTALHIYDKVMT 600
 DB 564 NPDVTQVGLVYVGSQVOTAFGLDTKPTRAAMLRAISOAPVYLGVGSAGTALHIYDKVMT 623
 QY 601 VQRGARPPGVKAVVVLTTGGRGAEDAAVPAQKIRNNGISVLYVGVGPVLSGIRLAPRD 660
 DB 624 VQRGARPPGVKAVVVLTTGGRGAEDAAVPAQKIRNNGISVLYVGVGPVLSGIRLAPRD 683
 QY 661 SLIHVAAYADLRHYDDVLIEMVLCGRAKOPVNLCKRSPCKNBSGCVLQNGSYRCKRDGWE 720
 DB 684 SLIHVAAYADLRHYDDVLIEMVLCGRAKOPVNLCKRSPCKNBSGCVLQNGSYRCKRDGWE 743
 QY 721 GPHCENRFLRRP 732
 DB 744 GPHCENRFLRRP 755
 RESULT 7
 ADNS9619
 ID ADNS9619 strand; protein; 755 AA.
 XX
 AC ADNS9619;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Colon neoplasia full-length ColoUp2 protein.
 XX
 KW molecular marker; neoplastic state; colon neoplasia; ColoUp1; ColoUp2;
 KW colon cancer.
 OS
 OS Unidentified.
 OS
 PN W02004018647-A2.
 XX
 PD 04-MAR-2004.
 XX
 PF 26-AUG-2003; 2003WO-US027085.
 XX
 PR 26-AUG-2002; 2002US-00229345.
 PR 18-OCT-2002; 2002US-00274177.
 XX
 PA (UYCA-) UNIV CASE WESTERN RESERVE.
 XX
 PI Markowitz SD;
 XX
 DR WPI; 2004-419536/39.
 XX
 PT Detecting likelihood of subject to have colon neoplasia, involves
 PT obtaining biological sample from subject, detecting one or more of
 PT ColoUp1 and ColoUp2 polypeptides, presence of polypeptide indicating
 PT colon neoplasia.
 XX
 PS Claim 11; SEQ ID NO 14; 118bp; English.
 XX
 CC The invention relates to novel molecular markers for categorizing the
 CC neoplastic state of a patient. The invention further comprises a method
 CC for detecting whether a subject is likely to have a colon neoplasia. The
 CC method comprises: obtaining a biological sample from the subject;
 CC detecting one or more polypeptides selected from among: one or more
 CC ColoUp1 polypeptides and one or more ColoUp2 polypeptides, wherein the
 CC presence of said one or more polypeptides is indicative of colon
 CC neoplasia. The invention further comprises a kit for carrying out the
 CC said method. The methods and compositions are useful for detecting colon
 CC cancer. This sequence represents a full-length ColoUp2 protein of the
 CC invention.

XX Sequence 755 AA;
 SQ Query Match 100.0%; Score 3826; DB 8; Length 755;
 Best Local Similarity 100.0%; Pred. No. 6,6e-300;
 Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LOEVHVSKEITGKISASRMMWCSAAVDIMFLDGSNSVGKGSFERSKHFATVCDGLDI 60
 24 LOEVHVSKEITGKISASRMMWCSAAVDIMFLDGSNSVGKGSFERSKHFATVCDGLDI 83

QY 61 SPERVAVGAFQFSSTPHLEFPPLDSFSTQGEVKARIKRMVFKGRTETELAKYLHRLGP 120
 DB 84 SPERVAVGAFQFSSTPHLEFPPLDSFSTQGEVKARIKRMVFKGRTETELAKYLHRLGP 143

QY 121 GGRNASVPOILLITVDKSGQDVALPSKQKRGVTVFVAVGRFPFWEELHALASPRQ 180
 DB 144 GGRNASVPOILLITVDKSGQDVALPSKQKRGVTVFVAVGRFPFWEELHALASPRQ 203

QY 181 HYLAAQVDAATNGLFSTLSSAICSSATPDCRVEAHPEHRTLEWVRFPAGNAPCMRGS 240
 DB 204 HYLAAQVDAATNGLFSTLSSAICSSATPDCRVEAHPEHRTLEWVRFPAGNAPCMRGS 263

QY 241 RRTLAVALAHCPEYSWKRVFLTHPATCYRTTCPCDSQPCQNGGTCVEBGLDGYQCLCP 300
 DB 264 RRTLAVALAHCPEYSWKRVFLTHPATCYRTTCPCDSQPCQNGGTCVEBGLDGYQCLCP 323

QY 301 LAFGBANCAKLSLECRVDLLFLDSSAGTTLDGFLRAKRVYKRFVRAVLSEDSRARG 360
 DB 324 LAFGBANCAKLSLECRVDLLFLDSSAGTTLDGFLRAKRVYKRFVRAVLSEDSRARG 383

QY 361 VATYSRELLVAVPVGVQVDPDLVMSLDGIPRPGCTPLTGSALROAERPGSATPTGD 420
 DB 384 VATYSRELLVAVPVGVQVDPDLVMSLDGIPRPGCTPLTGSALROAERPGSATPTGD 443

QY 421 RRRRVVLLTSHSEDEBAGAPAHARARELLLGVSSEAVRALEBITGSPKHWYSDP 480
 DB 444 RRRRVVLLTSHSEDEBAGAPAHARARELLLGVSSEAVRALEBITGSPKHWYSDP 503

QY 481 ODLFNOIPELOGLKCSRQRPGRCTQALDIVFLMILTASVGPENFAQMGSPVRSALQFV 540
 DB 504 ODLFNOIPELOGLKCSRQRPGRCTQALDIVFLMILTASVGPENFAQMGSPVRSALQFV 563

QY 541 NPDVTVGVVVGSSQVTAFGDLTKPTRAAMLPAISOAPVLGVSAGRALHTYKWT 600
 DB 564 NPDVTVGVVVGSSQVTAFGDLTKPTRAAMLPAISOAPVLGVSAGRALHTYKWT 623

QY 601 VORGARPGYKAVVLTGGRGADAAPPAQKLANNGISVLVVGVPVLSBGLRRLAGPRD 660
 DB 624 VORGARPGYKAVVLTGGRGADAAPPAQKLANNGISVLVVGVPVLSBGLRRLAGPRD 683

QY 661 SLIHVAAYADLRHYQDVLIELMLGSAKQPVNLCKPSPCMNMGSCVLQNGSYRCKRDWE 720
 DB 684 SLIHVAAYADLRHYQDVLIELMLGSAKQPVNLCKPSPCMNMGSCVLQNGSYRCKRDWE 743

QY 721 GPHCENRPLRRP 732
 DB 744 GPHCENRPLRRP 755

RESULT 8
 ADN39868
 ID ADN39868 standard; protein; 755 AA.
 XX
 AC ADN39868;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related polypeptide. SEQ ID NO: C238.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;

KW retinal neovascularization syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KW vulnerability; gene therapy; vaccine.
 XX Homo sapiens.
 XX WO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 XX 13-NOV-2002; 2002WO-US036810.
 XX
 XX 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359777P.
 PR 29-MAR-2002; 2002US-0368099P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397757P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevez PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 DR WPI: 2003-468649/44.
 XX
 DR N-PSDB; ADN39651.
 XX
 XX Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 12; SEQ ID NO C238; 1385pp; English.
 XX
 XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.
 XX
 XX
 SQ Sequence 755 AA;
 XX
 XX Query Match 99.9%; Score 3822; DB 7; Length 755;
 XX Best Local Similarity 99.9%; Pred. No. 1.4e-299;
 XX Matches 731; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITGKISASRMMWCSAAVDIMFLDGSNSVGKGSFERSKHFATVCDGLDI 60
 DB 24 LOEVHVSKEITGKISASRMMWCSAAVDIMFLDGSNSVGKGSFERSKHFATVCDGLDI 83

QY 61 SPERVRVGAFOFSSSTPHLEPPLDSFSTQGEVKARIKRMVFKGRTETELAKYLLHRGLP 120
 DB 84 SPERVRVGAFOFSSSTPHLEPPLDSFSTQGEVKARIKRMVFKGRTETELAKYLLHRGLP 143
 QY 121 GGRNASVPQILLIYVTOGKSQGDVALPSKOLKERGTVAVGVAPRMEELHALLASBPQ 180
 DB 144 GGRNASVPQILLIYVTOGKSQGDVALPSKOLKERGTVAVGVAPRMEELHALLASBPQ 203
 QY 181 HVLAEQVEDATNGLFSTLSSSAICSSATPDCEVAHPCERHTELVREBPAGNAPCMRGS 240
 DB 204 HVLAEQVEDATNGLFSTLSSSAICSSATPDCEVAHPCERHTELVREBPAGNAPCMRGS 263
 QY 241 RRTLAIVLAHCPFYSWKRVFLTHPATCYRTTCPCPCDSQPONGGTCVPEGLDGYQCLCP 300
 DB 264 RRTLAIVLAHCPFYSWKRVFLTHPATCYRTTCPCPCDSQPONGGTCVPEGLDGYQCLCP 323
 QY 301 LAFGEANCAKLTSLERVDLLFLDSSAGTTLDFLRAKVFVKRVAVLSEDSRAV 360
 DB 324 LAFGEANCAKLTSLERVDLLFLDSSAGTTLDFLRAKVFVKRVAVLSEDSRAV 383
 QY 361 VATYSRELLVAVPGEYQDVPDLVMSLDGIPFRGSPITLTSALRQAERGFSGATRTGOD 420
 DB 384 VATYSRELLVAVPGEYQDVPDLVMSLDGIPFRGSPITLTSALRQAERGFSGATRTGOD 443
 QY 421 RPRRVVLLTSHSEDEVAAPAHARARELLLLGVSEAVRAELEITGSPKRWVYSDP 480
 DB 444 RPRRVVLLTSHSEDEVAAPAHARARELLLLGVSEAVRAELEITGSPKRWVYSDP 503
 QY 481 QDFENQIPELQGLCSRORPGCTONLDVFMLDTSASVCPENFAOMOSFVRCALQFV 540
 DB 504 QDFENQIPELQGLCSRORPGCTONLDVFMLDTSASVCPENFAOMOSFVRCALQFV 563
 QY 541 NPDVTOGVLVVYSGQVOTAFGLDKPTBAAMLRAISQAPYLGVSAGTALLHIYDKWT 600
 DB 564 NPDVTOGVLVVYSGQVOTAFGLDKPTBAAMLRAISQAPYLGVSAGTALLHIYDKWT 623
 QY 601 VORGAPPGVPAVVLTGSRGAEDAAVPAQKLRNNGISVYLTVGVGVLSBGLRRLAGPRD 660
 DB 624 VORGAPPGVPAVVLTGSRGAEDAAVPAQKLRNNGISVYLTVGVGVLSBGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHNOVLTEMCGEAKOPNLCRSPCMWEGSCVQNGSYRCKCDGME 720
 DB 684 SLIHVAAYADLRHNOVLTEMCGEAKOPNLCRSPCMWEGSCVQNGSYRCKCDGME 743
 QY 721 GPHCENRFLRRP 732
 DB 744 GPHCENRFLRRP 755

RESULT 9
 ABG61891
 ID ABG61891 standard; protein; 807 AA.
 XX
 AC ABG61891;
 XX
 DT 15-AUG-2002 (first entry)
 XX
 DE Prostate cancer-associated protein #92.
 XX
 KM Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
 XX
 OS Mammalia.
 XX
 PN WO200230266-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 12-OCT-2001; 2001WO-US032045.
 XX
 PR 13-OCT-2000; 2000US-00687576.
 PR 08-DEC-2000; 2000US-00733288.
 PR 08-DEC-2000; 2000US-00733742.

PR 24-JAN-2001; 2001US-0263957P.
 PR 16-MAR-2001; 2001US-0276791P.
 PR 16-MAR-2001; 2001US-0276888P.
 PR 06-APR-2001; 2001US-0281922P.
 PR 24-APR-2001; 2001US-0286214P.
 PR 30-APR-2001; 2001US-00847046.
 PR 04-MAY-2001; 2001US-0288589P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Gish KC, Mack DH, Wilson KE, Afar D, Heyezi P;
 XX
 DR WPI; 2002-471335/50.
 DR N-PSDB; ABR22207.
 XX
 PT Detecting a prostate cancer-associated transcript in a cell in a patient,
 PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
 PT by determining if prostate cancer-associated genes are expressed in a
 PT prostate tissue.
 XX
 PS Claim 27; Page 377; 436pp; English.
 XX
 CC The present invention relates to methods of detecting a prostate cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with prostate cancer-
 CC associated polynucleotides (designated PC genes) that selectively
 CC hybridize to a sequence that is at least 80% identical to them. The
 CC prostate cancer-associated polynucleotide sequences are differentially
 CC expressed in prostate tumour tissue or in prostate cancer and are derived
 CC from the tissues of various organisms such as humans or other mammals
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
 CC associated genes are useful for diagnosing or treating prostate cancer,
 CC as well as for identifying modulators of prostate cancer or agents that
 CC inhibit prostate cancer. The nucleic acid sequences are particularly
 CC useful in gene therapy, as a vaccine or in antisense applications.
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins
 CC
 SQ Sequence 807 AA;
 XX
 Query Match 99.3%; Score 3799; DB 5; Length 807;
 Best Local Similarity 100.0%; Pred. No. 1,1e-297;
 Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LOEVNASKETIGKISAAKMMCSAAVDIMFLDSSNVGKGFERSKHFAITVCDGLDI 60
 DB 24 LOEVNASKETIGKISAAKMMCSAAVDIMFLDSSNVGKGFERSKHFAITVCDGLDI 83
 QY 61 SPERVRVGAFOFSSSTPHLEPPLDSFSTQGEVKARIKRMVFKGRTETELAKYLLHRGLP 120
 DB 84 SPERVRVGAFOFSSSTPHLEPPLDSFSTQGEVKARIKRMVFKGRTETELAKYLLHRGLP 143
 QY 121 GGRNASVPQILLIYVTOGKSQGDVALPSKOLKERGTVAVGVAPRMEELHALLASBPQ 180
 DB 144 GGRNASVPQILLIYVTOGKSQGDVALPSKOLKERGTVAVGVAPRMEELHALLASBPQ 203
 QY 181 HVLAEQVEDATNGLFSTLSSSAICSSATPDCEVAHPCERHTELVREBPAGNAPCMRGS 240
 DB 204 HVLAEQVEDATNGLFSTLSSSAICSSATPDCEVAHPCERHTELVREBPAGNAPCMRGS 263
 QY 241 RRTLAIVLAHCPFYSWKRVFLTHPATCYRTTCPCPCDSQPONGGTCVPEGLDGYQCLCP 300
 DB 264 RRTLAIVLAHCPFYSWKRVFLTHPATCYRTTCPCPCDSQPONGGTCVPEGLDGYQCLCP 323
 QY 301 LAFGEANCAKLTSLERVDLLFLDSSAGTTLDFLRAKVFVKRVAVLSEDSRAV 360
 DB 324 LAFGEANCAKLTSLERVDLLFLDSSAGTTLDFLRAKVFVKRVAVLSEDSRAV 383
 QY 361 VATYSRELLVAVPGEYQDVPDLVMSLDGIPFRGSPITLTSALRQAERGFSGATRTGOD 420
 DB 384 VATYSRELLVAVPGEYQDVPDLVMSLDGIPFRGSPITLTSALRQAERGFSGATRTGOD 443
 QY 421 RPRRVVLLTSHSEDEVAAPAHARARELLLLGVSEAVRAELEITGSPKRWVYSDP 480

```

Db      444 RRRRVVLLTSHSEDEVAGPARHARARELLLLGVSEAVRALEBITGSPKHWVWYSDP 503
Qy      481 QDLFNOIPELOGLKCSRRPGCCTQALDVFMLDTSASVGPENFAOMOSFVMSCALQFEV 540
Db      504 QDLFNOIPELOGLKCSRRPGCCTQALDVFMLDTSASVGPENFAOMOSFVMSCALQFEV 563
Qy      541 NPDVTOVGLVVGSGVQTAFGLDTKPTRAAMLRAISQAPYLGGVGSAGTALLHIYDKVMT 600
Db      564 NPDVTOVGLVVGSGVQTAFGLDTKPTRAAMLRAISQAPYLGGVGSAGTALLHIYDKVMT 623
Qy      601 VORGARPGVPKAVVVLTTGGRGAEDAAVPAQKLRNNGISVLVVGVPVLSBGLRRLAGPRD 660
Db      624 VORGARPGVPKAVVVLTTGGRGAEDAAVPAQKLRNNGISVLVVGVPVLSBGLRRLAGPRD 683
Qy      661 SLIHVAAYVADLRHYQDVLIEMLCGEAKOPVNLCKSPPCNNEGSCVYLQNGSYRCKCRDGE 720
Db      684 SLIHVAAYVADLRHYQDVLIEMLCGEAKOPVNLCKSPPCNNEGSCVYLQNGSYRCKCRDGE 743
Qy      721 GPHCENR 727
Db      744 GPHCENR 750

```

RESULT 10

ABP54687
ID ABP54687 standard; protein; 807 AA.

AC ABP54687;

DT 30-DEC-2002 (first entry)

DE Metastatic colorectal cancer-associated polypeptide.

KM Colorectal cancer; metastasis; differential expression; cytostatic;
diagnosis; gene therapy; vaccine.

OS Homo sapiens.

PN WO200268677-A2.

PD 06-SEP-2002.

PF 27-FEB-2002; 2002WO-US06001.

PR 27-FEB-2001; 2001US-0272206P.

PR 02-APR-2001; 2001US-0281149P.

PR 17-APR-2001; 2001US-0284555P.

PA (EOSB-) EOS BIOTECHNOLOGY INC.

PA (UYCA-) UNIV CASE WESTERN RESERVE.

PI Mack DH, Markowitz SD;

PI MPI; 2002-698677/75.

DR N-PSDB; AB081552.

XX New genes that are up- or down-regulated in colorectal cancer, useful for

XX diagnosing colorectal cancer in a subject, or for identifying modulators

XX of colorectal cancer-associated proteins and genes for treating

XX colorectal cancer.

XX Claim 8; Page 249; 260pp; English.

XX The present sequence is the protein sequence of a human polypeptide

XX encoded by a gene that exhibits increased expression in colon cancer-

XX derived metastases compared to normal colon tissue. It is an example of

XX claimed polypeptides that are encoded by genes which are differentially

XX expressed in metastatic colorectal cancer cells. Such polypeptides are

XX useful in diagnostic and prognostic assays, for raising antibodies useful

XX e.g. in immunotherapy, and in screening for modulator compounds of

XX therapeutic value

SQ Sequence 807 AA;

Query Match 99.3%; Score 3799; DB 5; Length 807;

Best Local Similarity 100.0%; Pred. No. 1.1e-297; Mismatches 0; Gaps 0;

Matches 727; Conservative 0; Indels 0; Gaps 0;

```

Qy      1 LOEVHVSKEETIKISAASKMMKMMCSAAVDMFLDSSNSYSGKS PERSKHAITVCDGLDI 60
Db      24 LOEVHVSKEETIKISAASKMMKMMCSAAVDMFLDSSNSYSGKS PERSKHAITVCDGLDI 83
Qy      61 SPERVVGAFOFSSPTPHLEPPLDSFSTQOEVARIKRMVFKCGRTETELAKYLLHRGLP 120
Db      84 SPERVVGAFOFSSPTPHLEPPLDSFSTQOEVARIKRMVFKCGRTETELAKYLLHRGLP 143
Qy      121 GGRNASVPOILLIIVDGSQGVVALPSYQIKERGTVFVGVPRFMEHLALASEPRQ 180
Db      144 GGRNASVPOILLIIVDGSQGVVALPSYQIKERGTVFVGVPRFMEHLALASEPRQ 203
Qy      181 HYLAEQVEDATNGLFSTLSSAICSSATPDCREAHPCERHRTLEWVEFAGNAPCMRGS 240
Db      204 HYLAEQVEDATNGLFSTLSSAICSSATPDCREAHPCERHRTLEWVEFAGNAPCMRGS 263
Qy      241 RRTLAIVLAHCPFPYSWKRVFLTHPATCYRTTCPCPCDSQPCONGTCVPEGIDGYOCLCP 300
Db      264 RRTLAIVLAHCPFPYSWKRVFLTHPATCYRTTCPCPCDSQPCONGTCVPEGIDGYOCLCP 323
Qy      301 LAFGEANALKLSIECRVDLLFLDSSAGTLLDGTLPKRVKX FVRVIVSEDSBARVG 360
Db      324 LAFGEANALKLSIECRVDLLFLDSSAGTLLDGTLPKRVKX FVRVIVSEDSBARVG 383
Qy      361 VATYSRELLVAVPGEYQDPDLVMSLDGI PRGGETLIGSALRQAAERFGSATRTQGD 420
Db      384 VATYSRELLVAVPGEYQDPDLVMSLDGI PRGGETLIGSALRQAAERFGSATRTQGD 443
Qy      421 RRRRVVLLTSHSEDEVAGPARHARARELLLLGVSEAVRALEBITGSPKHWVWYSDP 480
Db      444 RRRRVVLLTSHSEDEVAGPARHARARELLLLGVSEAVRALEBITGSPKHWVWYSDP 503
Qy      481 QDLFNOIPELOGLKCSRRPGCCTQALDVFMLDTSASVGPENFAOMOSFVMSCALQFEV 540
Db      504 QDLFNOIPELOGLKCSRRPGCCTQALDVFMLDTSASVGPENFAOMOSFVMSCALQFEV 563
Qy      541 NPDVTOVGLVVGSGVQTAFGLDTKPTRAAMLRAISQAPYLGGVGSAGTALLHIYDKVMT 600
Db      564 NPDVTOVGLVVGSGVQTAFGLDTKPTRAAMLRAISQAPYLGGVGSAGTALLHIYDKVMT 623
Qy      601 VORGARPGVPKAVVVLTTGGRGAEDAAVPAQKLRNNGISVLVVGVPVLSBGLRRLAGPRD 660
Db      624 VORGARPGVPKAVVVLTTGGRGAEDAAVPAQKLRNNGISVLVVGVPVLSBGLRRLAGPRD 683
Qy      661 SLIHVAAYVADLRHYQDVLIEMLCGEAKOPVNLCKSPPCNNEGSCVYLQNGSYRCKCRDGE 720
Db      684 SLIHVAAYVADLRHYQDVLIEMLCGEAKOPVNLCKSPPCNNEGSCVYLQNGSYRCKCRDGE 743
Qy      721 GPHCENR 727
Db      744 GPHCENR 750

```

RESULT 11

ABUS6622
ID ABUS6622 standard; protein; 807 AA.

AC ABUS6622;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polypeptide #215.

XX Lung cancer-associated polypeptide; cytostatic; emphysema;

XX antiinflammatory; antiaesthetic; non-small cell lung cancer; atelectasis;

XX small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

XX chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

XX

SEQ ID NO: 23

PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.

PS Claim 13; Page 304; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the proteins used
 CC for the detection method of the invention.

XX Sequence 807 AA;

Query Match 99.3%; Score 3799; DB 7; Length 807;
 Best Local Similarity 100.0%; Pred. No. 1.1e-297;
 Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEVHYSKRTIGTISAAKMMCSAAVDIMFLDGSNSYKGSFESKHPATVCGLDI 60
 DB 24 LOEVHYSKRTIGTISAAKMMCSAAVDIMFLDGSNSYKGSFESKHPATVCGLDI 83
 QY 61 SPERVVGAFAFSSSTHLEPPLDSFSTQOEVKARIKRMVFKGRTETELAKYLHRLGP 120
 DB 84 SPERVVGAFAFSSSTHLEPPLDSFSTQOEVKARIKRMVFKGRTETELAKYLHRLGP 143
 QY 121 GGRNAAVPOILIIITDQKSGDVALPSKQIKERGVTVPVAVGVPFPMEEIHALASSPQQ 180
 DB 144 GGRNAAVPOILIIITDQKSGDVALPSKQIKERGVTVPVAVGVPFPMEEIHALASSPQQ 203
 QY 181 HTLLAAQVEDATNGLTSTLSSAICSSATPDCVEAHPCGHRITLENVREPRAGNAPCWRS 240
 DB 204 HTLLAAQVEDATNGLTSTLSSAICSSATPDCVEAHPCGHRITLENVREPRAGNAPCWRS 263
 QY 241 RRTLLAVLAHCPFSYMKRVFLTPATCYRTTCGPPCDSPCCQCGTCVPEGLDGYCCLP 300
 DB 264 RRTLLAVLAHCPFSYMKRVFLTPATCYRTTCGPPCDSPCCQCGTCVPEGLDGYCCLP 323
 QY 301 LAFGEANCAKLKSLSCRVDLLFLDSSAGTTLDFLRAKVFYKRFVRAVLSHDSRAVY 360
 DB 324 LAFGEANCAKLKSLSCRVDLLFLDSSAGTTLDFLRAKVFYKRFVRAVLSHDSRAVY 383
 QY 361 VATYSRELLVAHVGEYQDVPDLVMSLDGIPFRGGPTLTGSALRQAAERFGSATRTGQD 420
 DB 384 VATYSRELLVAHVGEYQDVPDLVMSLDGIPFRGGPTLTGSALRQAAERFGSATRTGQD 443
 QY 421 RPRRVVULLTSHSEBEVGAHARARRELLLVGSEARALEERTSPKVMVYSDD 480
 DB 444 RPRRVVULLTSHSEBEVGAHARARRELLLVGSEARALEERTSPKVMVYSDD 503
 QY 481 QDFNQIPELQGLKCSRQPGCRTQALDLVFMLDTSASVPEPFAQMGSFVRSICALQFV 540
 DB 504 QDFNQIPELQGLKCSRQPGCRTQALDLVFMLDTSASVPEPFAQMGSFVRSICALQFV 563
 QY 541 NPVTOVGLVYVSSQVOTAFGLDTKPTRAAMLRAISQAPLVGGVGSAGTLLHITVYKWT 600
 DB 564 NPVTOVGLVYVSSQVOTAFGLDTKPTRAAMLRAISQAPLVGGVGSAGTLLHITVYKWT 623
 QY 601 VORGARPGVKAIVVLTGSGAGDAAPQAQLRNNGISVYVWGPVLSGLRLAGPRD 660
 DB 624 VORGARPGVKAIVVLTGSGAGDAAPQAQLRNNGISVYVWGPVLSGLRLAGPRD 683
 QY 661 SLIHVAAYADLRHYODVLEMLGSEAKQPNVLCPSQNNEGSCVLYONGSYRCKCDGME 720
 DB 684 SLIHVAAYADLRHYODVLEMLGSEAKQPNVLCPSQNNEGSCVLYONGSYRCKCDGME 743

QY 721 GPHCENR 727
 DB 744 GPHCENR 750

RESULT 13

ADN39869

ID ADN39869 strand1; protein; 807 AA.

XX ADN39869;

DT 17-JUN-2004 (first entry)

DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C239.

KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularization syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KW vulnery; gene therapy; vaccine.

XX Homo sapiens.

PN MO2003042661-A2.

PD 22-MAY-2003.

PF 13-NOV-2002; 2002WO-US036810.

PR 13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-0332464P.

PR 29-NOV-2001; 2001US-034393P.

PR 03-DEC-2001; 2001US-0335394P.

PR 14-DEC-2001; 2001US-0340376P.

PR 08-JAN-2002; 2002US-0347211P.

PR 10-JAN-2002; 2002US-0347349P.

PR 08-FEB-2002; 2002US-0355250P.

PR 13-FEB-2002; 2002US-0356714P.

PR 20-FEB-2002; 2002US-0359077P.

PR 29-MAR-2002; 2002US-036809P.

PR 04-APR-2002; 2002US-0370110P.

PR 12-APR-2002; 2002US-0372246P.

PR 05-JUN-2002; 2002US-0386614P.

PR 16-JUL-2002; 2002US-0396839P.

PR 22-JUL-2002; 2002US-0397775P.

PR 22-JUL-2002; 2002US-0397845P.

PR 09-SEP-2002; 2002US-0409450P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

PA Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX WPI, 2003-468649/44.

DR N-PSDB; ADN39862.

PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.

PS Claim 12; SEQ ID NO C239; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN3683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a

CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.

XX
 SO Sequence 807 AA;

Query Match 99.3%; Score 3799; DB 7; Length 807;
 Best Local Similarity 100.0%; Pred. No. 1.1e-297;
 Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LQEVHVKETIGTIGTISAAKMMKCSAAVDIMFLDGSNSVKGSPBSKHPATVCGDLDI 60
 DB 24 LQEVHVKETIGTIGTISAAKMMKCSAAVDIMFLDGSNSVKGSPBSKHPATVCGDLDI 83
 QY 61 SPBRVVGAFQFSSTPHLEPPLDSFSTQEVKARIKMWFKGRTETELAKYLHRLGP 120
 DB 84 SPBRVVGAFQFSSTPHLEPPLDSFSTQEVKARIKMWFKGRTETELAKYLHRLGP 143
 QY 121 GGNMNAVPOILLIIVTDGKSGDVALPSKOLKRGVTVFVAVGPRPMBELHALASEPRQ 180
 DB 144 GGNMNAVPOILLIIVTDGKSGDVALPSKOLKRGVTVFVAVGPRPMBELHALASEPRQ 203
 QY 181 HVLAAQVEDATNGTLFSTLSSAICSSATPDCEVAPHCNRLTENVREPAAGNAPCMRGS 240
 DB 204 HVLAAQVEDATNGTLFSTLSSAICSSATPDCEVAPHCNRLTENVREPAAGNAPCMRGS 263
 QY 241 RRTLAVALAHCIFYSMKRVFLTPATCYRTTCBPDCDOPCQNGCTVPEGLDGYQCLCP 300
 DB 264 RRTLAVALAHCIFYSMKRVFLTPATCYRTTCBPDCDOPCQNGCTVPEGLDGYQCLCP 323
 QY 301 LAAGGANCMLKSLERVDULLFLDSSACTTIDGFLRAVFKRVYRAVLSEDSRAVYG 360
 DB 324 LAAGGANCMLKSLERVDULLFLDSSACTTIDGFLRAVFKRVYRAVLSEDSRAVYG 383
 QY 361 VATYSRELLVAVVGVEYQDVPDLVWSLDGI PFRGSGPTLGSALROAAREGFGSATRTGDD 420
 DB 384 VATYSRELLVAVVGVEYQDVPDLVWSLDGI PFRGSGPTLGSALROAAREGFGSATRTGDD 443
 QY 421 RPRRVVLLTSHSEDEBAGAPAHARARELLLVGSEAVRALEBITGS PKHVMYSDP 480
 DB 444 RPRRVVLLTSHSEDEBAGAPAHARARELLLVGSEAVRALEBITGS PKHVMYSDP 503
 QY 481 QDLFNQIPBLQGLCSRRPGCTQALDVFMLDTSASVGPENFAQMQSFVASCALQFEV 540
 DB 504 QDLFNQIPBLQGLCSRRPGCTQALDVFMLDTSASVGPENFAQMQSFVASCALQFEV 563
 QY 541 NPDVTQGLVWVYSGOYQTAFGDLTKPTBAAMLRAISOAPLGGVGSAGTALHITYKWT 600
 DB 564 NPDVTQGLVWVYSGOYQTAFGDLTKPTBAAMLRAISOAPLGGVGSAGTALHITYKWT 623
 QY 601 VORGAPPGVPAVVVLTGGRGADAAVPAQKLRNNGISVLVVGVPVLSBGLRRLAGPRD 660
 DB 624 VORGAPPGVPAVVVLTGGRGADAAVPAQKLRNNGISVLVVGVPVLSBGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHGOVLIELMGEAKOPVNLCKSPCMNEGSCVLQNSYRCKCDGME 720
 DB 684 SLIHVAAYADLRHGOVLIELMGEAKOPVNLCKSPCMNEGSCVLQNSYRCKCDGME 743
 QY 721 GPHCENR 727
 DB 744 GPHCENR 750

RESULT 14
 ADN38804
 ID ADN38804 standard; protein; 807 AA.
 XX

AC ADN38804;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:122.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularization syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KW vulnery; gene therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN MO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 PP 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Gineburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 DR MPI: 2003-468649/44.
 XX
 DR N-PSDB; ADN38803.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 12; SEQ ID NO 122; 1385bp; English.
 XX
 CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.

SQ Sequence 807 AA;
 Query Match 99.3%; Score 3799; DB 7; Length 807;
 Best Local Similarity 100.0%; Pred. No. 1.1e-297;
 Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNSVGKSPERSKFAITVCDGLDI 60
 DB 24 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNSVGKSPERSKFAITVCDGLDI 83
 QY 61 SPERVVGAFOFSGSTHLEFPDLSFSTQGEVKARIKRWVFKGRTETELAKYLLHRLGP 120
 DB 84 SPERVVGAFOFSGSTHLEFPDLSFSTQGEVKARIKRWVFKGRTETELAKYLLHRLGP 143
 QY 121 GGRNAAVPOILLIYVTGKSGQDVALPSKQKRGTVFVAVGVPFPMWELHALLASPRQ 180
 DB 144 GGRNAAVPOILLIYVTGKSGQDVALPSKQKRGTVFVAVGVPFPMWELHALLASPRQ 203
 QY 181 HYLLAQVEDATNGLEFSTLSSSAICSSATPDCRVEAHPEBHRTLEHVRBPAAGNAPCMRGS 240
 DB 204 HYLLAQVEDATNGLEFSTLSSSAICSSATPDCRVEAHPEBHRTLEHVRBPAAGNAPCMRGS 263
 QY 241 RRTLAVLAHCEPFYSMKRVFLTHPATCYRTTCPCDSDPCQNGCTCPEGLDGYCLCP 300
 DB 264 RRTLAVLAHCEPFYSMKRVFLTHPATCYRTTCPCDSDPCQNGCTCPEGLDGYCLCP 323
 QY 301 LAFGGANCAKLSLECRVDLFLDSSAKTTIDGLRAKVPYKRVRAVLSDSRAARG 360
 DB 324 LAFGGANCAKLSLECRVDLFLDSSAKTTIDGLRAKVPYKRVRAVLSDSRAARG 383
 QY 361 VATYSRELLVAVGEYQDVPDLVMSLDGIPRKGFTLGSALROAAEGFGSATRTGDD 420
 DB 384 VATYSRELLVAVGEYQDVPDLVMSLDGIPRKGFTLGSALROAAEGFGSATRTGDD 443
 QY 421 RPRRVVLLTSHSEDEVAAPAHABARBEILLGVSEAVRALEBITGSPKIVMYSDP 480
 DB 444 RPRRVVLLTSHSEDEVAAPAHABARBEILLGVSEAVRALEBITGSPKIVMYSDP 503
 QY 481 ODIFNOIPELOGLCSRRORPGCTTOLLDVFMIDTNSAVGPFQMOGFWNSCALQFV 540
 DB 504 ODIFNOIPELOGLCSRRORPGCTTOLLDVFMIDTNSAVGPFQMOGFWNSCALQFV 563
 QY 541 NPDVTVGVLVVGSQVOTAFGLDTKPTRAMLEAIAOAPLGGVSGAGTALLHYIKWT 600
 DB 564 NPDVTVGVLVVGSQVOTAFGLDTKPTRAMLEAIAOAPLGGVSGAGTALLHYIKWT 623
 QY 601 VORGARPVKAVVLTGGRGADAAVPAQKLANNGISVLTGVGVLSBGLRRLAGPRD 660
 DB 624 VORGARPVKAVVLTGGRGADAAVPAQKLANNGISVLTGVGVLSBGLRRLAGPRD 683
 QY 661 SLIHVAAYADLRHYQVLIEMLCGEAKOPVNLCKPSPCMNEGSCVLTQNGSYRCKCRDGE 720
 DB 684 SLIHVAAYADLRHYQVLIEMLCGEAKOPVNLCKPSPCMNEGSCVLTQNGSYRCKCRDGE 743
 QY 721 GPHCENR 727
 DB 744 GPHCENR 750

RESULT 15
 ADN39446
 ID ADN39446 standard; protein; 807 AA.
 AC ADN39446;
 XX
 XX
 DT 17-JUN-2004 (first entry)
 XX
 XX
 DB Cancer/angiogenesis/fibrosis-related polypeptide. SEQ ID NO:A46.
 XX
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularization syndrome; scarring; uterine fibroid;

KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KW vulnery; gene therapy; vaccine.
 OS Homo sapiens.
 PN WO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002WO-US036810.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-035394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 DR WPI: 2003-468649/44.
 DR N-PDDB; ADN39445.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 12; SEQ ID NO A46; 1385bp; English.
 XX
 CC The invention relates to nucleic acid and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularization syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.
 XX
 SQ Sequence 807 AA;
 Query Match 99.3%; Score 3799; DB 7; Length 807;
 Best Local Similarity 100.0%; Pred. No. 1.1e-297;
 Matches 727; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNSVGKSPERSKFAITVCDGLDI 60
 DB 24 LOEVHVSKEITIGKISAAKMMWCSAAVDIMFLDGSNSVGKSPERSKFAITVCDGLDI 83

QY 61 SPERVRVGAFOFSSTPHLEFPDLSFSTQOEVKARIKRWVKGGRTELEALKYLLHRLGP 120
DB 84 SPERVRVGAFOFSSTPHLEFPDLSFSTQOEVKARIKRWVKGGRTELEALKYLLHRLGP 143
QY 121 GGNASVPOLLIIVTDGKSGDVALPSKOLKERGVTVPANVRPRMEELHALASEPRQ 180
DB 144 GGNASVPOLLIIVTDGKSGDVALPSKOLKERGVTVPANVRPRMEELHALASEPRQ 203
QY 181 HVLAAQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCHERTLBMVREFAGNAPCMRGS 240
DB 204 HVLAAQVEDATNGLFSTLSSSAICSSATPDCRVEAHPCHERTLBMVREFAGNAPCMRGS 263
QY 241 RRTLAVLAHCPRYSWKRVFLTHPATCYRTTCPCDCSOPCQNGTCVPEGLDGYOCLCP 300
DB 264 RRTLAVLAHCPRYSWKRVFLTHPATCYRTTCPCDCSOPCQNGTCVPEGLDGYOCLCP 323
QY 301 LAFGEANCAIKLSLECRVDLFLDSSAGTLDGFLRAKVFYKRFVRAVLSEDSRAVVG 360
DB 324 LAFGEANCAIKLSLECRVDLFLDSSAGTLDGFLRAKVFYKRFVRAVLSEDSRAVVG 383
QY 361 VATYSRELLVAVPVEGYQDVPDLVWSLDGI PFRGPTLTGSALROAERFGSATRTGOD 420
DB 384 VATYSRELLVAVPVEGYQDVPDLVWSLDGI PFRGPTLTGSALROAERFGSATRTGOD 443
QY 421 RPRRVVLLTSHSEDEPVAGPARHARARELLLGVGSEAVRALEBITGSPKIVMYSDP 480
DB 444 RPRRVVLLTSHSEDEPVAGPARHARARELLLGVGSEAVRALEBITGSPKIVMYSDP 503
QY 481 ODLFNQIPETQGLCSRORPGCRTOALDLVFMLDTSASVGPENFAOMOSFVRSALQFEV 540
DB 504 ODLFNQIPETQGLCSRORPGCRTOALDLVFMLDTSASVGPENFAOMOSFVRSALQFEV 563
QY 541 NPDVTQGLVWYGSQVQTAFGLDTKPTRAAMLRAISOAPYLGVGSGAGTALLHIYDKVMT 600
DB 564 NPDVTQGLVWYGSQVQTAFGLDTKPTRAAMLRAISOAPYLGVGSGAGTALLHIYDKVMT 623
QY 601 VQRGARPGVPKAVVLTGSRGABDAVPAQKLRNNGISVLVGVGPVLSBGLRRLAGPRD 660
DB 624 VQRGARPGVPKAVVLTGSRGABDAVPAQKLRNNGISVLVGVGPVLSBGLRRLAGPRD 683
QY 661 SLIHVAAYADLRHODVLIEMLCGEAKOPVNLCKSPCMNBGS CVLQNGSYRCKCRDGMW 720
DB 684 SLIHVAAYADLRHODVLIEMLCGEAKOPVNLCKSPCMNBGS CVLQNGSYRCKCRDGMW 743
QY 721 GPHCENR 727
DB 744 GPHCENR 750

Search completed: June 21, 2006, 20:07:57
Job time : 206 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OW protein - protein search, using sw model

Run on: June 21, 2006, 20:08:18 ; Search time 44 Seconds

(without alignments)
1600.697 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826

Sequence: 1 LQEVHVSKEITGKISASAKM.....CKCRDQWGPCHCENFLRRP 732

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	522.5	13.7	3176	2 CGH3A	collagen alpha 3(V
2	513.5	13.4	3137	2 A37797	collagen alpha 3(V
3	480.5	12.6	3124	2 A40020	collagen alpha 1(X
4	421.5	11.0	1857	2 S31312	collagen alpha 1(X
5	421.5	11.0	1888	2 S78476	collagen alpha 1(X
6	407.5	10.7	1747	2 A45974	collagen alpha 1(X
7	399	10.4	2813	1 VMH	von Willebrand fac
8	398	10.4	493	2 A33809	cartilage matrix p
9	390.5	10.2	500	2 S66522	cartilage matrix p
10	386.5	10.1	496	2 A37979	cartilage matrix p
11	305.5	8.0	741	2 T46488	hypothetical prote
12	301	7.9	2944	2 A54849	collagen alpha 1(V
13	290.5	7.6	550	2 T23760	hypothetical prote
14	288.5	7.5	3051	2 S42373	hypothetical prote
15	283	7.4	929	2 I51027	type XII collagen
16	265.5	7.0	567	2 T28797	hypothetical prote
17	266	7.0	843	2 A40970	undulin 1 - human
18	254	6.6	13055	2 T16580	hypothetical prote
19	225.5	5.9	1029	1 S21169	collagen alpha 2(V
20	224	5.9	1022	2 S04111	collagen alpha 2(V
21	216.5	5.7	724	2 A48569	antigen Em100 - B1
22	216.5	5.7	1018	1 CGH2A	collagen alpha 2(V
23	216	5.6	1019	1 A32856	collagen alpha 2(V
24	215	5.6	1025	2 S34839	collagen alpha 1(V
25	209.5	5.5	2098	2 T18397	protein CTR - mal
26	204.5	5.3	414	2 PS0323	von Willebrand fac
27	198.5	5.2	427	2 G00039	von Willebrand fac
28	192.5	5.0	371	2 S32604	collagen alpha 2(V
29	189.5	5.0	1523	2 T13953	MEGF5 protein - ra

30	189	4.9	1028	1 CGH3A	collagen alpha 1(V
31	188	4.9	1151	2 A45226	integrin alpha-1 c
32	184.5	4.8	712	2 A45638	immunodominant m/c
33	178.5	4.7	1180	2 A35854	integrin alpha-1 c
34	174.5	4.6	272	2 A55348	integrin alpha-1 -
35	172.5	4.5	2809	2 T30213	G-cadherin - sea u
36	172	4.5	1220	2 A6136	jagged protein pre
37	167	4.4	1328	2 T43060	agrin - electric r
38	166.5	4.4	1153	2 S00551	leukocyte surface
39	166.5	4.4	1163	1 RWHUIC	cell surface glyco
40	166	4.3	1955	1 AGCH	agrin precursor -
41	164.5	4.3	1286	2 A88396	protein MO1E10.2 l
42	164	4.3	640	2 T29784	hypothetical prote
43	164	4.3	2318	2 S45306	notch 3 protein -
44	164	4.3	2321	2 S78549	notch3 protein - h
45	162.5	4.2	746	2 S44792	P09G8.8 protein -

ALIGNMENTS

RESULT 1

CGH3A
collagen alpha 3(VI) chain precursor [validated] - human
N:Contains: collagen alpha 3(VI) chain, splice form A9/N10(-)
C:Species: Homo sapiens (man)
C:Date: 21-Nov-1993 #sequence revision 12-Nov-1999 #text change 16-Aug-2004
C:Accession: A59140; S13679; S24465; A57083; S28776; S00245; C31952; C29848; S26510; S48
R:Chu, M.L.
submitted to GenBank, May 1998
A:Reference number: A59140
A:Accession: A59140
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-3176 <CHU>
A:Cross-references: UNIPROT:P12111; UNIPARC:UPI000012604F; GB:XS2022; NID:G3127925; PIDN
R:Chu, M.L.; Zhang, R.Z.; Pan, T.; Stokes, D.; Conway, D.; Kuo, H.J.; Glanville, R.; May
EMBO J. 9, 385-393, 1990
A>Title: Mosaic structure of globular domains in the human type VI collagen alpha-3 chain
100b.
A:Reference number: S13679; MUID:90151612; PMID:1689238
A:Accession: S13679
A:Molecule type: mRNA
A:Residues: 1-30,237-313, 'CWM', 318-322, 'AR', 326-1815, 'FD', 1818-1819, 'ID', 1822-3176 <CH5>
A:Cross-references: UNIPARC:UPI0000173C14; EMBL:X52022; NID:G3127925
A:Accession: S24465
A:Molecule type: Protein
A:Residues: 574-585;965-973, 'X', 975-976;1306-1325;1361-1377;1381-1401;1473-1506, 'X', 1508
A:Residues: 1962, 'X', 1964-1965;2018-2037;2374-2410;2445-2459;2466-2469, 'X', 2471-2474;2504-2508, 'X',
A:Cross-references: UNIPARC:UPI0000173C15; UNIPARC:UPI0000173C16; UNIPARC:UPI0000173C17,
CIC; UNIPARC:UPI0000173C1D; UNIPARC:UPI0000173C1E; UNIPARC:UPI0000173C1F; UNIPARC:UPI000
10000173C25; UNIPARC:UPI0000173C26; UNIPARC:UPI0000173C27
A:Accession: S28776
A:Molecule type: DNA
A:Residues: 310-328 <ZAN>
A:Cross-references: UNIPARC:UPI0000173C28
A:Accession: S28776
A:Molecule type: mRNA
A:Residues: 32-126, 'AK', 129-136, 'L', 138-236 <ZAZ>
A:Cross-references: UNIPARC:UPI0000173C28; GB:IS49432; NID:G260296; PIDN:AA824261.1; PID
R:Chu, M.L.; Mann, K.; Deutzmann, R.; Pribula-Conway, D.; Hsu-Chen, C.C.; Bernard, M.P.,
Eur. J. Biochem. 168, 309-317, 1987
A>Title: Characterization of three constituent chains of collagen type VI by peptide se
A:Reference number: S00126; MUID:88029444; PMID:3665927
A:Accession: S00245
A:Molecule type: mRNA, protein
A:Residues: 2024-2046;2092-2156, 'R', 2203-2208, 'X', 2210-2211, 'X', 2213-2227;2228-2251;22314
A:Cross-references: UNIPARC:UPI000016A705; UNIPARC:UPI0000173C29; UNIPARC:UPI0000173C2A,
ID:G1335034

A>Note: the mRNA portion of the sequence corresponds to residues 2092-2157
 R.Chu, M.L.; Conway, D.; Pan, T.; Baldwin, C.; Mann, K.; Deutzmann, R.; Timpl, R.
 J. Biol. Chem. 263, 18601-18606, 1988
 A>Title: Amino acid sequence of the triple-helical domain of human collagen type VI.
 A:Reference number: A31952; MID:89066644; PMID:3198591
 A:Accession: C31952
 A:Molecule type: mRNA
 A:Residues: 2038-2373 <CH>
 A:Cross-references: UNIPARC:UPI0000173C2E; GB:J04211; GB:M20778
 A>Note: parts of this sequence were determined by protein sequencing
 R.Weil, D.; Matzel, M.G.; Passage, E.; Van Cong, N.; Pribula-Conway, D.; Mann, K.; Deutz
 Am. J. Hum. Genet. 42, 435-445, 1988
 A>Title: Cloning and chromosomal localization of human genes encoding the three chains of
 A:Reference number: A29848; MID:88161046; PMID:3348212
 A:Accession: C29848
 A:Molecule type: mRNA
 A:Residues: 2092-2151 <WEI>
 A:Cross-references: UNIPARC:UPI000016A720; GB:M27449; NID:g291919; PIRN:AAA52057.1; PID:
 A>Note: part of this sequence was determined by protein sequencing
 R.Jander, R.; Rautenberg, J.; Glanville, R.W.
 Eur. J. Biochem. 133, 39-46, 1983
 A>Title: Further characterization of the three polypeptide chains of bovine and human sh
 A:Reference number: S26506; MID:83209648; PMID:6852033
 A:Accession: S26510
 A:Molecule type: protein
 A:Residues: 'SALGVAGV' <VAN>
 A:Cross-references: UNIPARC:UPI0000173C2F
 A>Note: this sequence cannot be reliably placed and probably represents the results from
 R.Mayer, U.; Poeschl, E.; Nischt, R.; Specks, U.; Pan, T.C.; Chu, M.L.; Timpl, R.
 Eur. J. Biochem. 225, 573-580, 1994
 A>Title: Recombinant expression and properties of the Kunitz-type protease-inhibitor mod
 A:Reference number: S48709; MID:95045506; PMID:7525281
 A:Accession: S48709
 A:Molecule type: mRNA
 A:Residues: 'MRWIFPLCLAGRALA', 3102-3176 <MA>
 A:Cross-references: UNIPARC:UPI0000173C30
 A>Note: engineered sequence to allow isolation of the Kunitz-type proteinase inhibitor h
 R.Arnoux, B.; Meisigau, K.; Saludjian, P.; Norris, K.; Bjorn, S.; Olsen, O.;
 submitted to the Brookhaven Protein Data Bank, August 1994
 A:Reference number: A52812; PDB:1KNT
 A:Contents: annotation; X-ray crystallography, 1.6 angstroms, residues 3106-3160
 A>Note: engineered sequence expressed in *Saccharomyces cerevisiae* strain mt-663
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
 chains are 5-hydroxylated and subsequently O-glycosylated.
 C:Comment: The fibronectin type III repeat homology domain may be released during proces
 C:Gene: GDB:COL6A3
 A:Cross-references: GDB:119066; OMIM:120250
 A:Map position: 2q37.3-2q37.3
 C:Complex: type VI collagen is a heterotrimer of one alpha 1(VI) chain (see PIR:CGHUA),
 sociations: type trimer amino- and carboxyl-terminal domains (with disulfide bonds).
 C:Function:
 A:Description: structural component of extracellular tissue microfibrils associated with
 C:Keywords: alternative splicing; blocked amino end; cell binding; coiled coil; extracel
 F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-3176/Product: collagen alpha 3(VI) chain #status predicted <MAT1>
 F:26-3037/Domain: amino-terminal nonhelical #status predicted <AN>
 F:26-30-237-3176/Product: collagen alpha 3(VI) chain, splice form A9/N10(-) #status pred
 F:37-203/Domain: von Willebrand factor type A repeat homology <VM1>
 F:240-405/Domain: von Willebrand factor type A repeat homology <VM2>
 F:443-608/Domain: von Willebrand factor type A repeat homology <VM3>
 F:637-802/Domain: von Willebrand factor type A repeat homology <VM4>
 F:835-999/Domain: von Willebrand factor type A repeat homology <VM5>
 F:1027-1191/Domain: von Willebrand factor type A repeat homology <VM6>
 F:1231-1394/Domain: von Willebrand factor type A repeat homology <VM7>
 F:1434-1599/Domain: von Willebrand factor type A repeat homology <VM8>
 F:1637-1802/Domain: von Willebrand factor type A repeat homology <VM9>
 F:1835-2005/Domain: von Willebrand factor type A repeat homology <VM10>
 F:2038-2373/Region: von Willebrand factor type A repeat homology <VM10>
 F:2040-2042/Region: interrupted helical
 F:2135-2138/Region: cell attachment (R-G-D) motif
 F:2135-2150/Region: cell attachment (R-G-D) motif
 F:2148-2150/Region: cell attachment (R-G-D) motif

F:2154-2156/Region: cell attachment (R-G-D) motif
 F:2370-2372/Region: cell attachment (R-G-D) motif
 F:2374-3176/Domain: carboxyl-terminal nonhelical #status predicted <CNH>
 F:2400-2511/Domain: von Willebrand factor type A repeat homology <VM11>
 F:2612-2800/Domain: von Willebrand factor type A repeat homology <VM12>
 F:2865-2866/Region: von Willebrand factor type A repeat homology <VM12>
 F:2987-3072/Domain: fibronectin type III repeat homology <FN3>
 F:3111-3161/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>
 F:26/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
 F:108-116-202-251-2079-2311-2558-2677-2861-3036-binding site: carboxylate (Asn) (covale
 F:2087/Disulfide bonds: interchain #status predicted
 F:100-2206-2239-2316-2319/Modified site: 4-hydroxyproline (Pro) #status experimental
 F:2103-2209-2212-2322-2337-binding site: carboxylate (Lys) (covalent) #status experimental
 F:2103-2209-2212-2322-2337/Modified site: 5-hydroxylysine (Lys) #status experimental
 F:3111-3161-3120-3144-3136-3157/Disulfide bonds: #status predicted

Query Match 13.7% Score 532.5 DB 2: Length 3176;
 Best Local Similarity 25.5% Pred. No. 1.3e-28;
 Matches 166; Conservative 100; Mismatches 250; Indels 135; Gaps 17;

QY	25	AAVDIMPLDGSNSVGKSPERSKHFATVCCGLDISPERVVGAFQFSSTPHLEFPDLS	84
DB	36	AAADIIFLVDSWRTGEGHFQVREFLDYVKSALVANGENDRFPALVQNGNPHTEFLANT	95
QY	85	FSTQGEVAKIKRMVFKGRTETELAKYL--NRGLPGRNA--SVQIILIVTDKSKQ	140
DB	96	YRTKQEVLSHMSMYIGITQGTGKLEYIMOSHTRKAAGSPAGGVQVIVVLTGDSK	155
QY	141	GVVALPSQOKERGTVAVGVFRFMRBELALSEPFGQVYLA--QVEDANGKLS	197
DB	156	DGLAPSAELKADNVNPAIGEDABEALKEINSEPLNMHFNLENFTSLHDIQNLVS	215
QY	198	TLSSSAICSSATPPDCRVANHCENHRTLMVREFFAGNAPCMGSRRTTLVAACHCFYSWK	257
DB	216	CVHSSV-----SERRADTTLTDTNQ-----	238
QY	258	RVLTLTPATCYRTTCGPGCDSPCCNGSTCVPEGLDGYQCLPLAFGEBANCAKLSLEC	317
DB	239	-----DS-----	240
QY	318	RVDLLFLDSSAGTTLDFELAKVVRKFNVAVLSDBRAVGVATYRELLVAVPGEX	377
DB	241	ADIIFLDGSNNSTGVSNFVAIPLDLVNLKRLPIGTQOIRGVGVQFSDPTMFSLTQY	299
QY	378	QDVPLVMSLGIIPRGAPTTL--GSALRQAERGSATRRGQDPR---RVVLTRES	432
DB	300	STKQAVLAVKALGAGSELNLTGLDPTVVENF--TRAGSGRVEGSGQVAVLISAG	356
QY	433	HSEDEVAPARRHARRBELLLVGSSEAV--RALEBITGSPKRVVWVSDPQDLFNQIPELQ	491
DB	357	PSSDIRYGVVALKQASVFSFGLGAQASRAELQIATDDN--LVFTVPE--FRSFGDLQ	412
QY	492	GKLGCS-----RGRQAL-----DVFMLDTASVSPENFAMQSGVVRSCA	535
DB	413	EKLPLRYGVNAQRHLVKRPTITVQVLEVNKRDIPLVDSSALGLAFNMRIDTARVY	472
QY	536	LQFEVNPVTVGVLVVYSGVQTAFGLDTKPTRAMLRAT--SQAPYLGVGSAGTALLHI	594
DB	473	QRLTEGQDLIVNAQVADTVRPEFYTHHTTKKEVIRAVAKMKRDLDSALYTSALPQV	532
QY	595	YDKVMTVORG--ARRGVKAVVLTGSGAEDAAVPAQKLRRNGISVLVVG	643
DB	533	RNLTFTSSAGYRAAEGIPKLIVLTGKSLDEISQAPQLKRSSIMAPAIQ	583

RESULT 2
 A37797
 collagen alpha 3(VI) chain precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 16-Aug-2004
 C:Accession: A37797; A34270; A32674
 R:Poliana, R.; Bonaldo, F.; Colombatti, A.
 J. Cell Biol. 111, 2197-2205, 1990

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2006, 20:04:42 ; Search time 303 Seconds
(without alignment)
2234.691 Million cell updates/sec

Title: US-10-649-591-3
Perfect score: 3826
Sequence: 1 LQEVHVSKEITGKISASNM.....CKCRDQWGPCHCNFLRRP 732

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 7.2:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3826	100.0	755	05GFL6_HUMAN	05GFL6 homo sapien
2	3822	99.9	755	070U28_HUMAN	070U28 homo sapien
3	3540	92.5	725	06ZKJ7_HUMAN	06ZKJ7 homo sapien
4	3538	92.2	725	06ZS39_HUMAN	06ZS39 homo sapien
5	3135.5	82.0	791	070U27_MOUSE	070U27 mus musculu
6	3130.5	81.8	791	03UM88_MOUSE	03UM88 mus musculu
7	2901.5	75.8	721	08CE01_MOUSE	08CE01 mus musculu
8	2169.5	56.7	790	06DC06_XENLA	06DC06 xenopus lae
9	1766	46.2	717	04RG72_TETNG	04RG72 tetraodon n
10	551	14.4	1703	09Z019_MOUSE	09Z019 mus musculu
11	539.5	14.1	3169	04RKT6_CANFA	04RKT6 canis fam11
12	529.5	13.8	4421	04RKT6_TETNG	04RKT6 tetraodon n
13	522.5	13.7	2588	0530F4_HUMAN	0530F4 homo sapien
14	522.5	13.7	3176	06A333_HUMAN	06A333 homo sapien
15	514.5	13.4	647	04S2X7_TETNG	04S2X7 tetraodon n
16	513.5	13.4	3137	08C6K3_CHICK	08C6K3 gallus gall
17	510	13.3	1182	08C6K9_MOUSE	08C6K9 mus musculu
18	505	13.2	839	06P3N7_XENTR	06P3N7 xenopus tro
19	497.5	13.0	840	03ZNR2_XENLA	03ZNR2 xenopus lae
20	490.5	12.8	956	04RKT6_HUMAN	04RKT6 homo sapien
21	489	12.7	3119	05GFL6_MOUSE	05GFL6 mus musculu
22	484.5	12.7	937	06S6V0_MOUSE	06S6V0 mus musculu
23	484.5	12.7	956	099K64_MOUSE	099K64 mus musculu
24	483.5	12.6	928	03TQ80_MOUSE	03TQ80 mus musculu
25	483.5	12.6	956	08R542_MOUSE	08R542 mus musculu
26	483.5	12.6	955	08R542_MOUSE	08R542 mus musculu
27	480.5	12.6	955	05R5N1_PONPY	05R5N1 pongo pygma
28	480.5	12.6	1636	04TBC0_TETNG	04TBC0 tetraodon n
29	480.5	12.6	3124	04TBC0_TETNG	04TBC0 gallus gall
30	478.5	12.5	1796	05VYK2_HUMAN	05VYK2 homo sapien
31	475	12.4	2884	05VYK2_HUMAN	05VYK2 homo sapien

32	475	12.4	3063	1	COCA1_HUMAN	099715 homo sapien
33	475	12.4	3063	2	05VYK1_HUMAN	05VYK1 homo sapien
34	470	12.3	622	1	MATN4_HUMAN	095460 homo sapien
35	465	12.2	624	1	MATN4_MOUSE	089029 mus musculu
36	453	11.8	644	2	05NUJ1_BRARE	05nuj1 brachydanio
37	449.5	11.7	1222	2	063H04_HUMAN	063h04 homo sapien
38	449	11.7	726	2	05NUJ4_BRARE	05nuj4 brachydanio
39	447.5	11.7	1723	2	04SD22_TETNG	04sd22 tetraodon n
40	446.5	11.7	685	2	05NUJ2_BRARE	05nuj2 brachydanio
41	445.5	11.6	2657	2	088493_MOUSE	088493 mus musculu
42	444.5	11.6	1797	1	COBA1_MOUSE	080419 mus musculu
43	440	11.5	821	2	06PYX2_BRARE	06pyx2 brachydanio
44	439	11.5	950	2	04RKS9_TETNG	04rks9 tetraodon n
45	436.5	11.4	1259	2	04RP12_TETNG	04rp12 tetraodon n

ALIGNMENTS

RESULT 1

05GFL6_HUMAN PRELIMINARY; PRT; 755 AA.

AC 05GFL6; DT 01-MAR-2005, integrated into UniprotKB/TrEMBL.

DT 01-MAR-2005, sequence version 1.

DT 07-FEB-2006, entry version 6.

DE Colon cancer secreted protein-2.

GN Name=CCSP-2;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;

OC Homo.

OX NCBI_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX PubMed=15580307; DOI=10.1038/ef.0nc.1208134;

RA Xin B., Plutzer P., Fink S.P., Reese L., Nograti A., Willson J.K.,

RA Wilson K., Markowitz S.;

RT "Colon cancer secreted protein-2 (CCSP-2), a novel candidate

RT serological marker of colon neoplasia.";

RL Oncogene 24:724-731(2005).

CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms

CC Distributed under the Creative Commons Attribution-NonExiv License

CC EMBL; AY572973; AAT77226.1; -; Genomic_DNA.

DR EMBL; AY572972; AAT77225.1; -; mRNA.

DR Ensembl; ENSG00000165816; Homo sapiens.

GO; GO:0008083; F:growth factor activity; IEA.

DR InterPro; IPR006210; EGF.

DR InterPro; IPR000742; EGF.

DR InterPro; IPR001881; EGF_Ca_bd.

DR InterPro; IPR006209; EGF_like.

DR InterPro; IPR013032; EGF_like_reg.

DR InterPro; IPR002035; VWF_A.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00092; VWA; 3.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00181; EGF; 2.

DR SMART; SM00327; VWA; 3.

DR PROSITE; PS00022; EGF_1; UNKNOWN_1.

DR PROSITE; PS01186; EGF_3; 2.

DR PROSITE; PS00026; EGF_3; 2.

DR PROSITE; PS00234; VWF_A; 3.

KW EGF-like domain; Growth factor; Repeat.

SQ SEQUENCE 755 AA; 82012 MW; E02B99335B28BEC CRC64;

Query Match 100.0%; Score 3826; DB 2; Length 755;

Best Local Similarity 100.0%; Pred. No. 9.6e-26;

Matches 732; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LQEVHVSKEITGKISASNMCSAVIDIMFLDGSNVGKGSFERSKFAITVCDGLDI 60

```

Db      24  LQEVHVSKEITIGKISAASKMMWCSAAVDIMFLDGSNVGKGSFERSKHAITVCDGLDI 83
Qy      61  SPERVRVGAFOFSSSTPHLEFPDLSFTQOEVKARIKRMVFKGGRTETELAKYLLHRGIP 120
Db      84  SPERVRVGAFOFSSSTPHLEFPDLSFTQOEVKARIKRMVFKGGRTETELAKYLLHRGIP 143
Qy      121 GGNASVPOILLIIVTDGKSQGDVALPSKQKERGVTVFAVGVPFPMWEIHALASBPQQ 180
Db      144  GGNASVPOILLIIVTDGKSQGDVALPSKQKERGVTVFAVGVPFPMWEIHALASBPQQ 203
Qy      181  HVLAAQVEDATNGFLSTSSAISCATPDCRVEAHPCHEHRTLEWREFAAGNAPCWRS 240
Db      204  HVLAAQVEDATNGFLSTSSAISCATPDCRVEAHPCHEHRTLEWREFAAGNAPCWRS 263
Qy      241  RRTLAVALAACPFSYMKRVFLTPATCYRTTCGPGCDSPQCONGTCVPEGLDGYOCLCP 300
Db      264  RRTLAVALAACPFSYMKRVFLTPATCYRTTCGPGCDSPQCONGTCVPEGLDGYOCLCP 323
Qy      301  LAFGEANCAKLKLSLECRVDLFLDSSAGTTLDGFLRAKVPVKKFVRAVLSDBRARV 360
Db      324  LAFGEANCAKLKLSLECRVDLFLDSSAGTTLDGFLRAKVPVKKFVRAVLSDBRARV 383
Qy      361  VATYSRELLVAVPVGEGYDVPDLVMSLDGIPIFRGGPTLTGSALROAERGFSGATRTGQD 420
Db      384  VATYSRELLVAVPVGEGYDVPDLVMSLDGIPIFRGGPTLTGSALROAERGFSGATRTGQD 443
Qy      421  RPRRVVLLTSHSEDEVAQPARHARARELLLGVSSEAVRALEETITGSPKHVMYSDP 480
Db      444  RPRRVVLLTSHSEDEVAQPARHARARELLLGVSSEAVRALEETITGSPKHVMYSDP 503
Qy      481  ODLFNOIPELOGLKCSRORPGCTQALDVLVMDLTSASVGPENPAQOMOSFVNSCALQFEV 540
Db      504  ODLFNOIPELOGLKCSRORPGCTQALDVLVMDLTSASVGPENPAQOMOSFVNSCALQFEV 563
Qy      541  NPDVTQVGLVYVGSQVQTAFLGDTKPTRAAMLRAISOAPYLGGVSGAGTALLHYDKWMT 600
Db      564  NPDVTQVGLVYVGSQVQTAFLGDTKPTRAAMLRAISOAPYLGGVSGAGTALLHYDKWMT 623
Qy      601  VORGARPGVPKAVVYLTGGRGAEDAAVPAQKLRNNGISVLYVVGVPVLSBGLRRLAGPRD 660
Db      624  VORGARPGVPKAVVYLTGGRGAEDAAVPAQKLRNNGISVLYVVGVPVLSBGLRRLAGPRD 683
Qy      661  SLIHWAAVADLRHQVLIEMLCGEAKQPVNLCKPSPCMNBSGCVLQNGSYRCKCDGWE 720
Db      684  SLIHWAAVADLRHQVLIEMLCGEAKQPVNLCKPSPCMNBSGCVLQNGSYRCKCDGWE 743
Qy      721  GPHCENRFLRRP 732
Db      744  GPHCENRFLRRP 755

```

```

RESULT 2
Q70U28_HUMAN PRELIMINARY; PRT; 755 AA.
AC Q70U28;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 11.
DE AMACO precursor.
GN Name=AMACO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_Taxid=9606;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14506275; DOI=10.1074/jbc.M307794200;
RA Sengle G., Koble B., Morzelin M., Paulsson M., Wagener R.;
RT "Identification and characterization of AMACO, a new member of the von
RT Wilbrand factor A-like domain protein superfamily with a regulated
RT expression in the kidney.";
RL J. Biol. Chem. 278:50240-50249(2003).

```

```

CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial license
CC -----
DR EMBL; AJ536328; CAD60276.1; -, mRNA.
DR GO; GO:0008083; P.growth factor activity; IEA.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR000742; EGF_3.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR013032; EGF_Like_reg.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00092; VMA; 3.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00327; VMA; 3.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS50234; VMPA; 3.
KW EGF-like domain; Growth factor; Repeat; Signal.
FT SIGNAL 1
FT CHAIN 117..755 AMACO.
SQ SEQUENCE 755 AA; 82040 MW; 3144AACD34E31A4F CRC64;

Query Match
Best Local Similarity 99.9%; Score 3822; DB 2; Length 755;
Matches 731; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1  LQEVHVSKEITIGKISAASKMMWCSAAVDIMFLDGSNVGKGSFERSKHAITVCDGLDI 60
Db      24  LQEVHVSKEITIGKISAASKMMWCSAAVDIMFLDGSNVGKGSFERSKHAITVCDGLDI 83
Qy      61  SPERVRVGAFOFSSSTPHLEFPDLSFTQOEVKARIKRMVFKGGRTETELAKYLLHRGIP 120
Db      84  SPERVRVGAFOFSSSTPHLEFPDLSFTQOEVKARIKRMVFKGGRTETELAKYLLHRGIP 143
Qy      121 GGNASVPOILLIIVTDGKSQGDVALPSKQKERGVTVFAVGVPFPMWEIHALASBPQQ 180
Db      144  GGNASVPOILLIIVTDGKSQGDVALPSKQKERGVTVFAVGVPFPMWEIHALASBPQQ 203
Qy      181  HVLAAQVEDATNGFLSTSSAISCATPDCRVEAHPCHEHRTLEWREFAAGNAPCWRS 240
Db      204  HVLAAQVEDATNGFLSTSSAISCATPDCRVEAHPCHEHRTLEWREFAAGNAPCWRS 263
Qy      241  RRTLAVALAACPFSYMKRVFLTPATCYRTTCGPGCDSPQCONGTCVPEGLDGYOCLCP 300
Db      264  RRTLAVALAACPFSYMKRVFLTPATCYRTTCGPGCDSPQCONGTCVPEGLDGYOCLCP 323
Qy      301  LAFGEANCAKLKLSLECRVDLFLDSSAGTTLDGFLRAKVPVKKFVRAVLSDBRARV 360
Db      324  LAFGEANCAKLKLSLECRVDLFLDSSAGTTLDGFLRAKVPVKKFVRAVLSDBRARV 383
Qy      361  VATYSRELLVAVPVGEGYDVPDLVMSLDGIPIFRGGPTLTGSALROAERGFSGATRTGQD 420
Db      384  VATYSRELLVAVPVGEGYDVPDLVMSLDGIPIFRGGPTLTGSALROAERGFSGATRTGQD 443
Qy      421  RPRRVVLLTSHSEDEVAQPARHARARELLLGVSSEAVRALEETITGSPKHVMYSDP 480
Db      444  RPRRVVLLTSHSEDEVAQPARHARARELLLGVSSEAVRALEETITGSPKHVMYSDP 503
Qy      481  ODLFNOIPELOGLKCSRORPGCTQALDVLVMDLTSASVGPENPAQOMOSFVNSCALQFEV 540
Db      504  ODLFNOIPELOGLKCSRORPGCTQALDVLVMDLTSASVGPENPAQOMOSFVNSCALQFEV 563
Qy      541  NPDVTQVGLVYVGSQVQTAFLGDTKPTRAAMLRAISOAPYLGGVSGAGTALLHYDKWMT 600
Db      564  NPDVTQVGLVYVGSQVQTAFLGDTKPTRAAMLRAISOAPYLGGVSGAGTALLHYDKWMT 623
Qy      601  VORGARPGVPKAVVYLTGGRGAEDAAVPAQKLRNNGISVLYVVGVPVLSBGLRRLAGPRD 660

```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2006, 20:13:23 ; Search time 52 Seconds
(without alignments)
1232.162 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826

Sequence: 1 LQEVHVSKEITGKISASAKM.....CKCRDQWEGPHCENFLRRP 732

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:
1: /EMC Celerra_SIDS3/Ptodata/2/1aa/5_COMB.pep:*
2: /EMC Celerra_SIDS3/Ptodata/2/1aa/6_COMB.pep:*
3: /EMC Celerra_SIDS3/Ptodata/2/1aa/7_COMB.pep:*
4: /EMC Celerra_SIDS3/Ptodata/2/1aa/H_COMB.pep:*
5: /EMC Celerra_SIDS3/Ptodata/2/1aa/PCTUS_COMB.pep:*
6: /EMC Celerra_SIDS3/Ptodata/2/1aa/RE_COMB.pep:*
7: /EMC Celerra_SIDS3/Ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1580	41.3	405	2	US-09-312-283C-374 Sequence 374, App
2	499	13.0	915	2	US-09-907-794A-34 Sequence 34, Appl
3	499	13.0	915	2	US-09-905-125A-34 Sequence 34, Appl
4	499	13.0	915	2	US-09-902-775A-34 Sequence 34, Appl
5	499	13.0	915	2	US-09-906-700-34 Sequence 34, Appl
6	499	13.0	915	2	US-09-903-603A-34 Sequence 34, Appl
7	499	13.0	915	2	US-09-904-920A-34 Sequence 34, Appl
8	499	13.0	915	2	US-09-909-064-34 Sequence 34, Appl
9	499	13.0	915	2	US-09-905-381A-34 Sequence 34, Appl
10	499	13.0	915	2	US-09-906-618-34 Sequence 34, Appl
11	499	13.0	915	2	US-09-906-646-34 Sequence 34, Appl
12	499	13.0	915	2	US-09-904-462-34 Sequence 34, Appl
13	499	13.0	915	2	US-09-902-736A-34 Sequence 34, Appl
14	499	13.0	915	2	US-09-906-722A-34 Sequence 34, Appl
15	499	13.0	915	2	US-09-905-449-34 Sequence 34, Appl
16	499	13.0	915	2	US-09-905-562B-34 Sequence 34, Appl
17	499	13.0	915	2	US-09-906-679A-34 Sequence 34, Appl
18	499	13.0	915	3	US-09-907-841-34 Sequence 34, Appl
19	490.5	12.8	956	2	US-09-949-016-6215 Sequence 6215, Ap
20	490.5	12.8	963	2	US-09-949-016-11519 Sequence 11519, A
21	490.5	12.8	963	2	US-09-949-016-11520 Sequence 11520, A
22	484	12.7	959	3	US-10-000-512-10 Sequence 10, Appl
23	483.5	12.6	956	1	US-08-897-443-3 Sequence 3, Appl
24	477	12.5	584	2	US-09-949-016-10340 Sequence 10340, A
25	477	12.5	584	2	US-09-949-016-10341 Sequence 10341, A
26	451	11.8	1036	2	US-10-104-047-2812 Sequence 2812, Ap

27	426	11.1	2813	2	US-08-896-449A-2	Sequence 2, Appl
28	426	11.1	2813	2	US-09-132-652-2	Sequence 2, Appl
29	426	11.1	2813	2	US-09-886-900A-2	Sequence 2, Appl
30	426	11.1	2813	2	US-09-662-478C-2	Sequence 2, Appl
31	399	10.4	2050	1	US-08-347-594A-2	Sequence 2, Appl
32	396.5	10.4	2813	3	US-10-000-512-8	Sequence 8, Appl
33	396	10.4	2813	2	US-08-381-261A-1	Sequence 1, Appl
34	386.5	10.1	496	1	US-08-462-128-37	Sequence 37, Appl
35	386.5	10.1	496	1	US-08-463-180-37	Sequence 37, Appl
36	386.5	10.1	496	1	US-08-001-078A-1	Sequence 1, Appl
37	386.5	10.1	496	1	US-08-897-443-4	Sequence 4, Appl
38	386.5	10.1	496	1	US-08-463-218-1	Sequence 1, Appl
39	386.5	10.1	496	2	US-09-949-016-11306	Sequence 11306, A
40	386.5	10.1	496	5	PCT-US94-00253-1	Sequence 1, Appl
41	324.5	8.5	678	2	US-09-991-181-179	Sequence 179, App
42	324.5	8.5	678	2	US-09-990-444-179	Sequence 179, App
43	324.5	8.5	678	2	US-09-997-333-179	Sequence 179, App
44	324.5	8.5	678	2	US-09-992-598-179	Sequence 179, App
45	324.5	8.5	678	2	US-09-989-735-179	Sequence 179, App

ALIGNMENTS

RESULT 1									
US-09-312-283C-374									
Sequence 374, Application US/09312283C									
Patent No. 6573095									
GENERAL INFORMATION:									
APPLICANT: Watson, James D.									
APPLICANT: Strachan, Lorna									
APPLICANT: Sleeman, Matthew									
APPLICANT: Onrust, Rene									
APPLICANT: Murison, James G.									
TITLE OF INVENTION: Compositions isolated from skin cells									
FILE REFERENCE: 11000.1011c2									
CURRENT APPLICATION NUMBER: US/09/312,283C									
CURRENT FILING DATE: 1999-05-14									
NUMBER OF SEQ ID NOS: 425									
SOFTWARE: FaastSeq for Windows Version 4.0									
SEQ ID NO 374									
LENGTH: 405									
TYPE: PR									
ORGANISM: Mouse									
US-09-312-283C-374									
Query Match									
Best Local Similarity 78.6%; Pred. No. 1.8e-138;									
Matches 301; Conservative 29; Mismatches 49; Indels 4; Gaps 2;									
QY	1	LQEVHVSKEITGKISASAKM	MMCSAAVDIMFLDSSNSVKGSPFRSKHFAITVCGLDI	60					
DB	24	LQEVHVSKEITGKISASAKM	MMCSAAVDIMFLDSSNSVKGSPFRSKHFAITVCGLDI	83					
QY	61	SEPRRVGAFQSSPHLEFLDSSNSVKGSPFRSKHFAITVCGLDI	120						
DB	84	SEPRRVGAFQSSPHLEFLDSSNSVKGSPFRSKHFAITVCGLDI	142						
QY	121	GGRNASVPDILITVDGSKQGVAPLPSKOLKRGVTVFVAVGRFPRMEELHALLASEPRQ	180						
DB	143	GGRNASVPDILITVDGSKQGVAPLPSKOLKRGVTVFVAVGRFPRMEELHALLASEPRQ	202						
QY	181	HVLAEOVEDATNGLFSTLSSAICSSATPDGCVZAHPCHEHTLMEVAFAGNAPCWRS	240						
DB	203	HVLAEOVEDATNGLFSTLSSAICSSATPDGCVZAHPCHEHTLMEVAFAGNAPCWRS	262						
QY	241	RTTLAVLAHAGFYSWKVFLTHPATCYRTTTPRGCDSPQPCNGGTCVPEGDQCCP	300						
DB	263	RTTLAVLAHAGFYSWKVFLTHPATCYRTTTPRGCDSPQPCNGGTCVPEGDQCCP	322						
QY	301	LAFGGANCAKLTLSCEKVDLFLDSSAGTTLDFLRAKVFVKFVAVLSEDSRARVG	360						

Db 323 LAGGEVNCAPKSLBERIDVLFLLSSAGTTLGGFRRAKAFVKRFVQAVLRDSDRAVQ 382
Qy 361 VATYSRELVAVPVGEYQDVPDL 383
Db 383 IASGRMLWAVAVC---RGVPAL 402

RESULT 2
US-09-907-794A-34
; Sequence 34, Application US/0990794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltgen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gunney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavir, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Thomas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 34
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-907-794A-34

Query Match 13.0%; Score 499; DB 2; Length 915;
Best Local Similarity 23.9%; Pred. No. 4e-37;
Matches 188; Conservative 101; Mismatches 264; Indels 232; Gaps 28;

Qy 28 DIMPLDGSNSVGKSPERSKHFAITVCDDGISPERVGAFOFSPHLEFPDPSFT 87
Db 57 DVFIIIDSSRSVNTHDYAKVEFYVDIIQFLDIGDVTFRVGLQGSTVKNFEISLKTFR 116
Qy 88 QOEVARIKRVFKGRTETELAKYLLHRGLPGGRNA---SVPOLIIVTDKSGQD 142
Db 117 KSEVERAVRMRHLSTGMTGLAIQYALINAFSEABGARPLRENPVIMITDGRPODS 176
Qy 143 VALPSKOLKEGCVTVFANGVFRPMEELHALASERGGYVLLA---EYVEDATNGLPSTL 199
Db 177 VAEVAKAKADTGILFALIGVQVDPTLTKISGEHEDHVFVAVNFSQIETLTSVFOKLT 236
Qy 200 SSSAICSSATPD-----CRVHAHPCHEHRTLE 225
Db 237 CTAMKCTLEHNCAMFCINIRGSYCRCKQYIILSDOTTCTRIQDLCAEDHNCQLCVN 296
Qy 226 MYREBAGNAPCWRG-----SRRTLAVALAHCP-----FYSWRRVFLT 262
Db 297 VPGSFV--CQCYSGYALAEGRCAVAV--DYCASBNHCEHCVNADGSYLCQCHEGFAL 352
Qy 263 HP--ATCYRTT-C---PGPDSQ-----PCQNGCTVCPGL----- 292
Db 353 NDEKTCRIRYCALANKPG-CEHECVNNEESYCHRGYTLDPNGKTCRVDHCAQDDH 411
Qy 293 -----DOYGLCPDLAF-----GGEANCAKLJLEGRVLDL 322
Db 412 GGEQLCLNTBDSFVQCSEGFILNEDLTKCSRVYDCLSDHCEYSCV-----NMDS 464
Qy 333 FLLDSSAGTTL--DGFRLAKV-----FV-KRFRVAVLSEDSRARV 359
Db 465 FACQCPBGHVLRSQDKTCAKLDSCALGDHGEHSCVSESDSFVCCFGYILRED----- 519
Qy 360 GVATYSRELVAVPVGEYQDVPDLVMSLDGIPFRGGPPTLTGSALRQAAREGGSATRTGQ 419
Db 520 GKTCCRKQVCAIDHG-----CEHICVNSDSDTCECT-----EGF---RLAE 559
Qy 420 D--RPRRVVLLTSHSEDEVA---GPARHARARELLLVGSEAVRAELEBITSPKHV 474
Db 560 DKKRRKRDVCKSTHGCHEHICVANNNGSYICKCSGFVL-----AE----- 600
Qy 475 MYVSDPDLFNOIPELQKLCRSRORPGCRTOALDLVFMULTSASVGPENFAMQSFVNSC 534
Db 601 -----DGRCKK-----CTEGRIDIVFYDSSKSGENFEVVKQFVGI 640
Qy 535 ALQFVNPDVTVQGLVTVGSGVQVAFGLDTPTRAAMLRAISQAFYLGGVGSAGTALHI 594
Db 641 IDSLTISPKARVGLQYSTQVHTEFTLRNFSADMKKAAVAHMKYMGKSGMTGIALKGM 700
Qy 595 YDKVTVQGRAR---GVPKAVVVLTCGRGADEDAVAPQKLRNNGISVLYGVGVFLSEG 651
Db 701 FERSSTQSGARPLSTVRPRAIVTDSRADDDVSEMAKKAANGITMYAVGVGAIEBE 760
Qy 652 LRLRLA 656
Db 761 LQELIA 765

RESULT 3

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2006, 20:25:52 ; Search time 23 Seconds

(without alignments)
727.085 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826

Sequence: 1 LQEVNVSKEITGKISASNM.....CKCRDGEHPHCENFLRRP 732

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 99297 seqs, 2284552 residues

Total number of hits satisfying chosen parameters: 99297

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.New:*
1: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEM_PUB.pep.*
2: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEM_PUB.pep.*
3: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US07_NEM_PUB.pep.*
4: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US08_NEM_PUB.pep.*
5: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US09_NEM_PUB.pep.*
6: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US10_NEM_PUB.pep.*
7: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US11_NEM_PUB.pep.*
8: /EMC_Celerra_SIDS3/ptodata/1/pubpaa/US06_NEM_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	324.5	8.5	678	US-10-196-749-150	Sequence 150, App
2	324.5	8.5	678	US-11-101-316-34	Sequence 34, App1
3	285	7.4	550	US-10-196-749-46	Sequence 46, App1
4	242	6.3	180	US-09-970-076-3	Sequence 27, App1
5	210	5.5	769	US-11-165-586-27	Sequence 592, App
6	189	4.9	436	US-10-505-928-592	Sequence 103, App
7	188	4.9	1151	US-11-246-999-103	Sequence 45, App1
8	184.5	4.8	724	US-11-165-586-45	Sequence 36, App1
9	179.5	4.7	765	US-11-165-586-36	Sequence 831, App1
10	178	4.7	2026	US-10-505-928-831	Sequence 21, App1
11	173.5	4.5	1218	US-11-178-724-21	Sequence 290, App
12	173.5	4.5	1523	US-10-196-749-290	Sequence 162, App
13	164.5	4.3	1167	US-11-175-714-162	Sequence 22, App1
14	163.5	4.3	1228	US-11-178-724-22	Sequence 168, App
15	158.5	4.1	1198	US-09-970-076-4	Sequence 165, App
16	157.5	4.1	1166	US-11-175-714-166	Sequence 165, App
17	157.5	4.1	1188	US-11-175-714-165	Sequence 38, App1
18	157.5	4.1	1188	US-10-196-749-38	Sequence 15, App1
19	157	4.1	737	US-11-296-092-15	Sequence 43, App1
20	157	4.1	737	US-11-296-092-15	Sequence 148, App
21	157	4.1	737	US-11-296-155-15	Sequence 147, App
22	155.5	4.1	1034	US-11-246-999-43	
23	155.5	4.1	1141	US-11-175-714-150	
24	155.5	4.1	1166	US-11-175-714-148	
25	155.5	4.1	1188	US-11-175-714-147	

26	155.5	4.1	1189	US-11-246-999-35	Sequence 35, App1
27	153	4.0	1152	US-10-511-937-2444	Sequence 2444, App
28	150.5	3.9	437	US-11-293-697-3386	Sequence 3386, App
29	147.5	3.9	182	US-11-175-714-159	Sequence 159, App
30	145.5	3.8	182	US-11-175-714-151	Sequence 151, App
31	134.5	3.5	247	US-11-330-353-4	Sequence 4, App1
32	130.5	3.4	696	US-10-449-902-41608	Sequence 41608, A
33	128.5	3.4	1170	US-10-511-937-3007	Sequence 3007, App
34	126	3.3	4391	US-11-183-325-56	Sequence 56, App1
35	125.5	3.3	618	US-11-178-724-19	Sequence 19, App1
36	125.5	3.3	1047	US-11-293-697-3852	Sequence 3852, App
37	124.5	3.3	536	US-10-449-902-40167	Sequence 40167, A
38	122.5	3.2	685	US-11-175-714-2	Sequence 2, App1
39	122.5	3.2	685	US-11-175-714-28	Sequence 28, App1
40	122.5	3.2	685	US-11-175-714-32	Sequence 32, App1
41	122.5	3.2	685	US-11-175-714-34	Sequence 34, App1
42	122.5	3.2	685	US-11-175-714-36	Sequence 36, App1
43	122.5	3.2	685	US-11-175-714-38	Sequence 38, App1
44	122.5	3.2	685	US-11-178-724-20	Sequence 20, App1
45	122.5	3.2	830	US-11-175-714-11	Sequence 11, App1

ALIGNMENTS

RESULT 1
US-10-196-749-150
Sequence 150, Application US/10196749
GENERAL INFORMATION:
Publication No. US20060094864A1
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C340
CURRENT APPLICATION NUMBER: US/10/196,749
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 150
LENGTH: 678
TYPE: PRT
ORGANISM: Homo Sapien
US-10-196-749-150

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 22, 2006, 12:17:24 ; Search time 10858 Seconds
(without alignments)
6466.593 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826

Sequence: 1 LGEVHVSKEITGKISASIAK.....CKCRDGEQPHCNRLRNP 732

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ .p2n.model -DEV=xlh
-Qc/abs/ABSSMBE.epool/US10649591/runat.21062006.161710.8168/app_query.fasta.1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs807
-USER=US10649591.QCEN_1.1.5548.Qrunat.21062006.161710.8168 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-NARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_env.*
2: gb_pat.*
3: gb_ph.*
4: gb_pl.*
5: gb_dr.*
6: gb_pro.*
7: gb_ro.*
8: gb_sts.*
9: gb_gy.*
10: gb_un.*
11: gb_vl.*
12: gb_ov.*
13: gb_hcg.*
14: gb_in.*
15: gb_cm.*
15: gb_da.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3826	100.0	3097	5	AY572972 Homo sapi
2	3822	99.9	2647	5	AF536328 Homo sapi
3	3540	92.5	2946	2	C0842058 Sequence

4	3540	92.5	2946	5	AK122716	AK122716 Homo sapi
5	3528	92.2	2782	2	C0850594	C0850594 Sequence
6	3528	92.2	2782	5	AK127756	AK127756 Homo sapi
7	3135.5	82.0	2653	6	MM0536329	MM0536329 Mus muscu
8	2194.5	57.4	1837	2	BD209726	BD209726 Compositi
9	2194.5	57.4	1837	2	AR341532	AR341532 Sequence
10	2169.5	56.7	3777	11	BC077945	BC077945 Xenopus
11	1958.5	51.2	1220	2	C0728687	C0728687 Sequence
12	1616.5	42.3	52487	5	AY572973	AY572973 Homo sapi
13	1616.5	42.3	121774	5	AC005383	AC005383 Homo sapi
14	1477.5	38.6	154390	6	AC125150	AC125150 Mus muscu
15	1477.5	38.6	176356	6	AC161518	AC161518 Mus muscu
16	1375.5	36.0	235745	12	AC119647	AC119647 Rattus no
17	787	20.6	161750	11	CR352283	CR352283 Zebrafish
18	787	20.6	222348	12	CR931999	CR931999 Danto rer
19	786.5	20.6	194152	11	BX65198	BX65198 Zebrafish
20	786.5	20.6	194444	12	CR932360	CR932360 Danto rer
21	551	14.4	5321	6	AF034136	AF034136 Mus muscu
22	529	13.8	8667	2	C0720672	C0720672 Sequence
23	525	13.7	10427	7	BV175219	BV175219 sqm77639
24	525	13.7	10488	2	CS080598	CS080598 Sequence
25	525	13.7	10537	7	BV175220	BV175220 sqm77640
26	525	13.7	10538	2	C0833988	C0833988 Sequence
27	525	13.7	10558	2	CQ874659	CQ874659 Sequence
28	525	13.7	10558	2	CQ981613	CQ981613 Sequence
29	525	13.7	10558	2	CS030819	CS030819 Sequence
30	525	13.7	10558	2	CS039771	CS039771 Sequence
31	525	13.7	10558	2	AX393287	AX393287 Sequence
32	525	13.7	10558	5	HS020287	HS020287 H. sapiens R
33	525	13.7	10558	2	CQ412877	CQ412877 Sequence
34	519	13.6	2756	2	AX675291	AX675291 Sequence
35	518.5	13.6	2361	2	C0716197	C0716197 Sequence
36	513.5	13.4	10057	11	CH000441	CH000441 Chicken alp
37	509.5	13.3	3550	5	BC010444	BC010444 Homo sapi
38	508.5	13.3	3711	2	C093664	C093664 Sequence
39	507.5	13.3	3711	2	CQ714974	CQ714974 Sequence
40	505	13.2	2943	11	BC063920	BC063920 Xenopus t
41	505	13.2	3373	2	AR085071	AR085071 Sequence
42	504	13.2	3449	2	BD075394	BD075394 Secretory
43	504	13.2	3449	2	BD172254	BD172254 Secreted
44	504	13.2	3449	2	BD172573	BD172573 Secreted
45	504	13.2	3449	2	BD172892	BD172892 Secreted

ALIGNMENTS

RESULT 1
AY572972
LOCUS
DEFINITION
Homo sapiens colon cancer secreted protein-2 (CCSP-2) mRNA, complete cds.
ACCESSION
AY572972
VERSION
AY572972.1
KEYWORDS
GI:50429311
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 3097)
Xin,B., Platzzer,P., Pink,S.P., Reese,L., Nosrati,A., Willson,J.K., Wilson,K. and Markowitz,S.
Colon cancer secreted protein-2 (CCSP-2), a novel candidate serological marker of colon neoplasia
Oncogene 24 (4), 724-731 (2005)
JOURNAL
PUBMED
15580307
REFERENCE
2 (bases 1 to 3097)
Xin,B., Platzzer,P., Pink,S.P., Nosrati,A., Reese,L.M., Willson,J.K., Wilson,K. and Markowitz,S.D.
Direct Submision
Submitted (13-MAR-2004) Medicine, Case Western Reserve University, 11001 Cedar Ave., Cleveland, OH 44106, USA
LOCATION/Qualifiers

```

source
1. .3097
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q25-q26"
1. .3097
/gene="CCSP-2"
89. .2356
/gene="CCSP-2"
/ncbi="colon cancer diagnostic marker"
/codon_start=1
/product="colon cancer secreted protein-2"
/protein_id="AAT77225.1"
/db_xref="GI:50429312"
/translation="MPPFLLEAVCVLPSPVPSPLOEVHVKETIGKISAKMM
WCSAAVDIMFLDGSNSVGKSPERSNHRATYCDGLDIPBVRVGAPOPSSTPHLE
PFLDPSFQOEYKARIKMYFKGRTETELATKLHRLGGKMSVPTLITVDG
KSGQDVALPSKQLEKGVFAVGVRPRHEILALASEPRGHVLALEDATNGL
FSTLSSSAICSAIPDCRVABPCHEHTELEMBEFAGNAPCMRSGRTTAVLAHCP
YSWKRVLTPATCYRTCPGCDSPCCQNGCTVPEGLDGYCLCPLEAGBANAL
KLSECRVDLFLDSSAGTTLDPLRAKVFVRPRAVISEDRAVGATVSRRL
VAVPGEYODVPDPLVMSLDGIPRGGPTLTGSAIROAABRBSATRTGDRRARIV
LUTESHSEDEVAPARARARLELLAGVSEAVRABLEITGSPKVMVTSDDLPN
QILELQKLCSSRQPCGRTQALDVFELDTASVGBENFAQMSFVRSCLQPEVND
VTQVGLVVGSGVQTAFGLDTKPTRAMLRALISQAPLGGVSGATRALHITDKMTV
ORGARGPVPAVVVLITGSRGAEDAAVPAQCLRNNGISVLVVGVPVLSERLAPR
DSLHYAAVADLRHODVLEMLGSEAKQPVNLCKPSPCMNBSCVLQNSYVCKCRD
GMEGPHCEHFLRFP"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 0 Length: 3097
Score: 3826.00 Matches: 732
Percent Similarity: 100.0% Conservative: 0
Beet Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

```

US-10-649-591-3 (1-732) x AY572972 (1-3097)

```

QY 1 LeuGlnGluValHisValSerIyRgluThrIleGlyValIleSerAlaIAserIyMet 20
DB 158 CTCGAGAAAGTCCATGTAAGCAAGAAACCATCGGGAAGATTTCAGCTGCACAAATG 217
QY 21 MetTrpCysSerAlaValAspIleMetPheLeuLeuAspGlySerAsnSerValGly 40
DB 218 ATGTGGTCTCGGCTGCAGTGCATCATATTTCTGTAGATGGGCTTAACAGCGTCGGG 277
QY 41 LysGlySerPheGluIuArgSerIyHisPheAlaIleThrValCysAspGlyLeuAspIle 60
DB 278 AAGGAGAGCTTTGAAAGGTCCAAAGCACTTGGCCATCACTGTGACGGTTCGACATC 337
QY 61 SerProGluIuArgValArgValIleGlyIlePheGlnPheSerSerThrProHisLeuGluPhe 80
DB 338 AGCCCGAGAGGGTGCAGAGTGGAGCAATTCAGTTCACTCCATCTCGAAGATTC 397
QY 81 ProLeuAspSerPheSerThrGlnGlnGluValIleValIleArgIleIySarMetValPhe 100
DB 398 CCTTGATTCATTTTCAACCCCAAGAGTGAAGGCAGAGATCAAGAGAGATGTTTTTC 457
QY 101 LysGlyValArgThrGluThrgluLeuAlaLeuIySerIleLeuHisIArgGlyLeuPro 120
DB 458 AAGGAGGGGCGACGAGGAGCGAATCTTGTAATTACTTCTGCACAGAGGTTGGCT 517
QY 121 GlyGlyArgAsnAlaSerValProGlnIleLeuIleIleValThrAspGlyIySerGln 140
DB 518 GAGGAGGAATCTCTCTGTGCCCCAGATCTCATCATCTCATGGAGAGTCCAG 577
QY 141 GlyAspValAlaLeuProSerIyGlnLeuIyGlyIuArgGlyValIleThrValPheAlaVal 160
DB 578 GGGGATGTGCACTGCATCCAAAGCACTGAAGAAAGGGGTGTCACTGTGTTGCTGTG 637

```

```

QY 161 GlyValArgPheProArgTrpGluGluLeuHisValIleLeuAlaSerGluProArgGlyGln 180
DB 638 GGGGTCAAGTTTCCAGGTGGAGGAGGTGCATGCACTGGCCAGCCGCTTAAGGGCAG 697
QY 181 HisValLeuLeuAlaGluGlnValGluIuAspAlaThrAsnGlyLeuPheSerThrLeuSer 200
DB 698 CAGTCTGTGGCTTGAAGAGGTGAGAGATGCCAACAGGCTTTCAGCACCTCAGC 757
QY 201 SerSerAlaIleCysSerSerAlaThrProAspCysArgValGluAlaHisProCysGlu 220
DB 758 AGCTCGGCACTGTGCTCCAGCCGCCAGCCAGACTGCAGGGTCAAGGCTCAACCTTGA 817
QY 221 HisArgThrLeuGluMetValArgIuPheAlaGlyAsnAlaProCysTrpArgIySer 240
DB 818 CACAGAGACCTGAGATGGTCCGGAGGTTCCGTGGCAATGCCCATGTCTGGAAGATGG 877
QY 241 ArgArgThrLeuAlaValLeuAlaAlaHisCysProPheIySerTrpIyAspValPhe 260
DB 878 CGCGGACCCCTGCGGTGTGCTGCACACTGTCTTACAGCTGGAAAGAGATGTTC 937
QY 261 LeuThrHisProAlaThrCysTrpArgTrpThrCysProGlyProCysAspSerGlnPro 280
DB 938 CTAAACCACCTGCCACCTGCTACAGACCACTGCCAGGCCCTGTGACTGCAGCCC 997
QY 281 CysGlnAsnGlyValYThrCysValProGluGlyLeuAspGlyTrpGlnCysLeuCysPro 300
DB 998 TGCAGAAATGAGAGCAGATGTGTCCAGAAAGACTGAGCGGCTACCAAGTGCCTCCGCC 1057
QY 301 LeuAlaPheGlyGlyGluAlaAsnCysValAlaLeuIyLeuSerLeuGluCysArgValAsp 320
DB 1058 CTGGCTTTGGAGGGAGGCTAACTGTGCTCCGAACTGAAGCTGAATGCAAGGCTCAGC 1117
QY 321 LeuLeuPheLeuLeuAspSerSerAlaGlyThrThrLeuAspGlyPheLeuArgAlaIy 340
DB 1118 CTCTCTTCTCTGCTGCAGACGCTCTGCGGCACCACTGAGCGCTTCTTGGCGCCAA 1177
QY 341 ValPheValIyArgPheValArgValAlaLeuSerGluAspSerArgAlaArgValGly 360
DB 1178 GTCTTGTAAGCGGTTGTGCGGGCCGTGTGAGGAGAGACTCTCGGGCCGAGTGGGT 1237
QY 361 ValAlaThrTrpSerArgGluLeuLeuValAlaValProValGlyIuTrpGlnAspVal 380
DB 1238 GTGGCCACATACAGCAGGAGGCTGTGTGGGTGCTGTGGGAGAGTACAGAGATGG 1297
QY 381 ProAspLeuValTrpSerLeuAspGlyIleProPheArgGlyValProThrLeuThrGly 400
DB 1298 CTGACCTGTGTGAGAGCTTCATGGCAATCTCTTCCGTGTGGTCCCACTGACGGGC 1357
QY 401 SerAlaLeuArgGlnAlaAlaGluIuArgGlyPheGlySerAlaIleThrArgThrGlyGlnAsp 420
DB 1358 AGTGCTTGCAGGCGGCAAGCGGCAAGCGTGGCTTCGGGAGGCCACAGACAGGCCAGGAC 1417
QY 421 ArgProArgArgValValValLeuLeuThrGluSerHisSerGluAspGluValAlaGly 440
DB 1418 CGGCACTGTAAGTGTGTGTTTGTCTCACTGAGTCACTCGAGAGATAGGTTGGGGCC 1477
QY 441 ProAlaArgHisAlaArgAlaArgGluLeuLeuLeuGlyValGlySerGlnAlaVal 460
DB 1478 CAGGCCGCTCAGCAGAGGGCGGAGAGCTGCTCTCTGTGGGTGAGGCACTGAGCGCG 1537
QY 461 ArgAlaGluLeuGluGluIleThrGlySerProIyHisValMetValIySerAspPro 480
DB 1538 CGGGCAGAGCTGAGAGAGATCACAGGAGCCCAAGCATGTATGTTACTCGAGTCT 1597
QY 481 GlnAspLeuPheAsnGlnIleProGluLeuGlnGlyIyLeuCysSerArgGlnAspPro 500
DB 1598 CAGATCTGTTAAACCAATCTCTAGCTGCAGGGGAAAGCTGTGACGCGGAGCGGCA 1657
QY 501 GlyCysArgThrGlnAlaLeuAspLeuValPheMetLeuAspThrSerAlaSerValGly 520
DB 1658 GGGTCCGAGACCAAGCCCTGACCTGTCTTCACTTGTGACACCTCTCCCTCAGTAA 1717
QY 521 ProGluAsnPheAlaGlnMetGlnSerPheValArgSerCysAlaLeuGlnPheGluVal 540

```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2006, 12:16:14 ; Search time 116 Seconds
(without alignments)
6680.217 Million cell updates/sec

Title: US-10-649-591-3
Perfect score: 3826
Sequence: 1 LQEVHVKETIGKISASKM.....CKCRDGMGEGHCENRFLRRP 732

Scoring table:	
BLOSUM62	
Xgapop 10.0 ,	Xgapext 0.5
Ygapop 10.0 ,	Ygapext 0.5
Fgapop 6.0 ,	Fgapext 7.0
Delop 6.0 ,	Delext 7.0

Searched: 5244920 seqs, 3486124231 residues
Total number of hits satisfying chosen parameters: 10489840

```

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

```

Command line parameters:
-MODEL=frama+ p2n. model -DEV=xjh
-O=/abs/ABBSWEB/aspot/US10649591/runat_21062006.161708.8157/app_query.fasta
-DB=/abs/ABBSWEB/aspot/US10649591/runat_21062006.161708.8157/app_query.fasta
-UNITs=bits STRAP=1 -END=1 -MATH=tblcnm62 -TRANS=humano cd1 -LIST=5
-DOCLIGN=200 THR_SCORE=pct THR_MAX=100 THR_MIN=0 -ALIGN=5 -MODE=LOCAL
-OUTFMT=pct -NORM=none -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=dsb07
-USER=US10649591_QCGN_1_1_761 -ORIGIN_21062006.161708.8157 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORE=0 -WAIT -DSBLOCK=2100 -LONGLOG -DEV_TIMESTAMP=120
-WARN_TIMESTAMP=30 -THEBADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGADOP=6 -FGAPEXT=7
-XGADOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database :
1_ GeneSeq_8:*
2:   geneSeqn1980s:*
3:   geneSeqn1990s:*
4:   geneSeqn2000s:*
5:   geneSeqn2001s:*
6:   geneSeqn2002s:*
7:   geneSeqn2002bs:*
8:   geneSeqn2003as:*
9:   geneSeqn2003bs:*
10:  geneSeqn2003cs:*
11:  geneSeqn2003ds:*
12:  geneSeqn2004as:*
13:  geneSeqn2004bs:*
14:  geneSeqn2005s:*
15:  geneSeqn2006s:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3826	100.0	2810	8	ABX76352 Abx76352 Lung cancer
2	3826	100.0	2810	10	ADB80522 Ovarian cancer
3	3826	100.0	2810	11	ADN38805 Cancer/astrocytoma

4	3826	100.0	281.0	12	AD170218	Ad170218 Colon cancer
5	3826	100.0	281.0	12	ADN59620	ADN59620 Colon neo
6	3822	99.9	2268	11	ADN39651	ADn39651 Cancer/an
7	3799	99.3	2424	6	ABQ81552	Abq81552 Gene up-r
8	3799	99.3	2424	8	ABX76351	Abx76351 Lung canc
9	3799	99.3	2424	10	ADB80520	ADb80520 Ovarian c
10	3799	99.3	2424	11	ADN39652	ADn39652 Cancer/an
11	3799	99.3	2424	11	ADN38803	ADn38803 Cancer/an
12	3799	99.3	3375	6	ABK92207	Abk92207 Prostate
13	3799	99.3	3375	11	ADN39445	ADn39445 Cancer/an
14	3799	99.3	3375	11	ADN39458	ADn39458 Cancer/an
15	3799	99.3	3375	11	ADN39564	ADn39564 Cancer/an
16	3799	99.3	3375	12	ADQ70145	AdQ70145 Human COL
17	3540	92.5	2946	12	ADQ63544	AdQ63544 Novel hum
18	3528	92.5	2782	13	ADG75557	ADg75557 DNA leng
19	3387.5	88.4	3485	13	ASB81530	ASb81530 DFN encod
20	3382.5	88.4	2481	8	ABX76350	Abx76350 Lung canc
21	3382.5	88.4	2481	11	ADN80518	ADb80518 Ovarian c
22	3382.5	88.4	2481	11	ADN38801	ADn38801 Cancer/an
23	3178.5	83.1	5464	12	ADG78853	ADg78853 Human can
24	3178.5	83.1	5543	12	ADG78852	ADg78852 Human can
25	2194.5	57.4	1837	3	AAZ61803	Aaz61803 CDNA enc
26	2194.5	57.4	1837	4	ACG99736	ACg99736 Skin cell
27	2194.5	57.4	1837	6	AB134888	Ab134888 Murine cf
28	927	24.2	352	12	ADQ21815	AdQ21815 Human gsf
29	530.5	13.9	9075	13	ACN41218	ACn41218 Human dia
30	529	13.8	10542	13	ACN37863	ACn37863 Tumour-as
31	525	13.7	10057	13	ADP54236	ADp54236 Human PRO
32	525	13.7	10488	14	AD222646	AD222646 Collagen
33	525	13.7	10556	6	AA159459	AA159459 Human pol
34	525	13.7	10558	6	AB192097	Ab192097 Human Tum
35	525	13.7	10558	8	ABX04155	ABx04155 Human mRN
36	525	13.7	10558	8	ABZ76357	ABz76357 Human alp
37	525	13.7	10558	10	ABX72022	ABx72022 DNA encod
38	525	13.7	10558	11	ADP65610	ADp65610 Human RNA
39	525	13.7	10558	12	ADP65057	ADp65057 Human gsf
40	525	13.7	10558	12	ADQ18246	AdQ18246 Human co5
41	525	13.7	10558	12	ADQ29558	ADq29558 Human COL
42	525	13.7	10558	13	ADG87615	ADg87615 Human TYR
43	525	13.7	10558	13	ADU06244	ADu06244 Novel btr
44	525	13.7	10558	14	ADV70142	ADv70142 Tumour-as
45	525	13.7	10558	14	ADY14519	ADy14519 DNA encod

ALIGNMENTS

RESULT 1
ABX76352
ID ABX76352 standard; DNA; 2810 BP.

AC ABX76352;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polynucleotide #216.

KM Lung cancer-associated polynucleotide; gene; de cytosatic; emphysema;
KM antiinflammatory; antiaesthetic; non-small cell lung cancer; atelectasis
KM small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KM chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KM interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

OS Unidentified

PN WO200286443-A2.

PD 31-OCT-2002.

PF 18-APR-2002; 2002WO-US012476.

PR 18-APR-2001; 2001US-0284770P.
PR 10-MAY-2001; 2001US-0290492P.
PR 09-NOV-2001; 2001US-0339245P.

PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 PI Aziz N, Murray R;
 DR MPI; 2003-093161/08.
 XX P-PSDB; ABUS6623.
 PT
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 for treating lung cancer, by contacting a biological sample from the
 patient with a polynucleotide that exhibits increased or decreased
 expression in lung cancer.
 PT
 PS Claim 22; Page 351-352; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
 CC invention
 XX
 SQ Sequence 2810 BP; 554 A; 801 C; 880 G; 575 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 3, 93e-244 Length: 2810
 Score: 3826.00 Matches: 732
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-649-591-3 (1-732) x ABX76352 (1-2810)
 QY 1 LeuGInGUvAlhIeVAlSeRlYsGJuThrIlleGlyLeVleIeSeRAlaASeRlYsMet 20
 Db 158 CTCcAGGAAGTCATCTTAAGCAAGAAACCATCGGAAAGTTTCACTCCACGCAAAAG 217
 QY 21 MetTrpCySeSerAlaAlaValApIlIeMetPheLeuLeuApGlySeRAnSeRValGly 40
 Db 218 ATGTGGTGTCCGCTCGCTCAGTGCATCATGTTTCTGTAGATGGCTTAACAGCGTCGG 277
 QY 41 LysGlySeRPhesGluRSeRlYsHIsPheAlaIlleThrValCySapGlyLeuApIlIe 60
 Db 278 AAAGGGAGCTTTAAAGGTCCAAGCACTTGGCATCAAGTCGTACAGGCTCGGACATC 337
 QY 61 SerProGluuArgValaArgValaIaphesGlnPheSeSerThRProHIsleuGluPhe 80
 Db 338 AGCCCGAAGAGGGTCAAGTGGAGCATTCAGTTCCACTCTCATCTGGAATTC 397
 QY 81 ProLeuApSeRPheserThRgInGInGUvAlIyAlaArgIlIeYsArgMetValPhe 100
 Db 398 CCTTGATTCATTTCACCCCAACGAGAAAGTGAAGGCAAGATCAAGAGATGCTTTC 457
 QY 101 LysGlyGlyArgThRgInThRgInUleuAlaLeuLYeRlYleuLeuHIsArgGlyLeuPro 120
 Db 458 AAAGAGGGGGCGACGGAAGACGGAACCTTGCTGAAATACCTTCTGCAAGAGGGTTCCT 517

QY 121 GlyGlyArgAsnAlaSeRValProGInUleuIlleLeuAlIeValThRAspGlyYsSerGln 140
 Db 518 GGAAGGAGAAATGCTTCTGTGCCCCAGATCCATCATCATGTCATCATGATGGAAAGTCCAG 577
 QY 141 GlyAspValAlaLeuProSeRlYsGlnLeuLYsGlyValaIlyThValPheAlaVal 160
 Db 578 GGGAGATGGGCACTGCTCCATCCAAAGCACTGAAGAAAGGGGTGTCACTGTGTTTGTGTG 637
 QY 161 GlyValArgPheProArgTrpGluGluLeuHIsAlaLeuAlaSeRGlupProArgGlyGln 180
 Db 638 GGGGTCAAGTTTCCAGAGTGGAGAGCTGCATGCATCTGGCCAGGAGCTTGAAGGCGAG 697
 QY 181 HisValLeuLeuAlaGluGlnValaIyAspAlaThRAsnGlyLeuPheSeRThRLeuSeR 200
 Db 698 CACGTGCTGTGCTGAGCAGGTGAGGATGCCAACACAGGCTCTTCAGACACCTCAGC 757
 QY 201 SerSerAlaIleCySeSerSerAlaThRProApCySArgValaIyAlaIleProCySglu 220
 Db 758 AGCTGGCCATCTGCTCCAGGCCACGCCAGCACTCAGAGGTGAGAGCTCACCCCTGTAG 817
 QY 221 HisArgThRLeuGluMetValaArgGluPheAlaGlyAsnAlaProCyETrpArgGlySer 240
 Db 818 CACAGGACCTGGAATGCTCCGGAGTTTCGTGGCATGCCCCATGCTGGAGAGGATG 877
 QY 241 ArgArgThRLeuAlaValleuAlaAlaHIsCySProPheTrYSerTrpYsArgValaPhe 260
 Db 878 CGGCGGACCTTGCGGTGTGCTGCACTGCTCCCTTCAAGCTGGAAAGAGAGTTC 937
 QY 261 LeuThRlAProAlaThRcySlyrArgThRThRcySProGlyProCyAspSerGlnPro 280
 Db 938 CTAAACCAACCCCTGCACTGCTACAGAGCACTCCAGGCCCTGTGATCTGCGAGGCC 997
 QY 281 CysGlnAsnGlyGlyThRcySValaProGluGlyLeuAspGlyTYrGlnCySLeuCySPro 300
 Db 998 TCCCAAGATGAGGCAATGTGTTCCAGAAAGACTGAGCGCTTACAGTGCCTTGCCTCG 1057
 QY 301 LeuAlaPheGlyGlyGluAlaAsnCySAlaLeuLYsLeuSerLeuGluCySArgValaAsp 320
 Db 1058 CTGGCTTTGGAGGGAGGCTAACTGTGCCCTGAAGCTGAGCTGGAATGAGAGGTCGAC 1117
 QY 321 LeuLeuPheLeuLeuApSeSerAlaGlyThRThRleuAspGlyPheLeuArgAlaYs 340
 Db 1118 CTCCTCTTCTGCTGTGACAGCTCTGCGGACACCTCTGGAAGGCTCTTCTGCGGCCAA 1177
 QY 341 ValPheValIyAsArgPheValaArgAlaValleuSerGluAspSerArgAlaArgValaGly 360
 Db 1178 GTCTTCGTGAAGCGGTTTGTGCGGGCCGTGCTGACGAGGACTTCTCGGGCCGAGTGGGT 1237
 QY 361 ValAlaThRTrYSerArgGluLeuLeuValaIyAlaProValaGlyGluTYrGlnAspVal 380
 Db 1238 GTGGCCACATACAGCAGGAGAGCTGCTGTGGGGTGCCTGTGGGGAGTACAGAGATGG 1297
 QY 381 ProAspLeuValaIlyPheSerleuAspGlyIleProPheArgGlyGlyProThRleuThRgIy 400
 Db 1298 CCTGACCTGTGCTGAGCCTCATGATGCATCTTCTCGTGTGGGCCCAACCTTGACGGGG 1357
 QY 401 SerAlaLeuArgGlnAlaAlaGluArgGlyPheGlySerAlaThRArgThRgIyGlnAsp 420
 Db 1358 AGTGCCTTGGCGCAGGCGGAGAGCTGCTTCCGGAAGCGCCACAGAGAGGCGGAGAC 1417
 QY 421 ArgProArgArgValaValaIleuLeuThRgInUSeRHisSerGluAspGluValaIyGly 440
 Db 1418 CGGCCACGTAAGAGTGTGGTTTGTCTCACTAGTCACACTCCGAGGATGAGGTTGGGGCG 1477
 QY 441 ProAlaArgHIsAlaArgAlaArgGluLeuLeuLeuLeuGlyValaGlySerGluAlaVala 460
 Db 1478 CCAAGCGGTCAACGAAGGGCGGAGAGCTGCTCCGTGCGGTGTGAGGAGTGAAGCCGG 1537
 QY 461 ArgAlaGluLeuGluGluIlyethRgLYSeRProLYsHIsValMetValTYrSerAspPro 480
 Db 1538 CGGGCAGAGCTGGAGAGATCAAGGCAAGCCCAAGCATGTGATGTCTACTCGAATCCT 1597

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 22, 2006, 12:21:53 ; Search time 8668 Seconds
(without alignments)
7083.459 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826
Sequence: 1 LGVHVSKETIKISASAKM.....CKCRDQWEGPHCENFLRRP 732

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues
Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=xlh
-Q=/abs/ABSWEB.epool/US1064591/runat_21062006.161713.8205/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-OCCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs04
-USER=US1064591@cscn_1_1.6323@runat_21062006.161713.8205 -NCPD=6 -ICPD=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-RAN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_est7:*
7: gb_est8:*
8: gb_est9:*
9: gb_est10:*
10: gb_est11:*
11: gb_est12:*
12: gb_est13:*
13: gb_est14:*
14: gb_est15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3135.5	82.0	3216	6 AK133486	Mus muscu
2	3130.5	81.8	3039	6 AK145058	Mus muscu
3	3117.5	81.5	2948	6 AK166339	Mus muscu
4	3109.5	81.3	3888	6 AK160971	Mus muscu

5	3109	81.3	3627	6 AK029297	Mus muscu
6	3034.5	79.3	2395	14 DQ050604	DO050604 Homo sapi
7	2162	56.5	2261	14 DQ050605	DO050605 Pan trogl
8	1234	32.3	896	10 D7808453	D7808453 LB0174.CR
9	1003	26.2	585	9 DA191273	DA191273 DA191273
10	866.5	22.6	1024	7 BF164945	BF164945 601778179
11	772.5	20.2	632	11 AZ344444	AZ344444 1M0078P24
12	706.5	18.5	664	7 BB617396	BB617396 BB617396
13	688	18.0	801	9 CX921133	CX921133 JGI CAAN8
14	677	17.7	577	9 DB146766	DB146766 DB146766
15	654	17.1	469	5 CK392905	CK392905 K0851D11-
16	652	17.0	851	3 BU271772	BU271772 D03506709
17	626	16.4	374	4 BX956774	BX956774 DKF2P781D
18	621	16.2	744	5 CF285997	CF285997 AGENCOURT
19	610.5	16.0	638	7 BB613951	BB613951 BB613951
20	556	14.5	804	10 DT201883	DT201883 JGI CAAS1
21	554.5	14.5	572	3 BU262922	BU262922 603816005
22	553.5	14.5	461	9 DN334113	DN334113 LIB185-0
23	541.5	14.2	420	7 BB846685	BB846685 BB846685
24	529	13.8	9534	14 DQ052789	DQ052789 Homo sapi
25	510	13.3	3955	6 AK054356	AK054356 Mus muscu
26	510	13.3	4728	6 AK087828	AK087828 Mus muscu
27	509.5	13.3	2760	14 AY402756	AY402756 Homo sapi
28	508	13.3	529	9 DA791716	DA791716 DA791716
29	499	13.0	9184	14 AY402246	AY402246 Homo sapi
30	497.5	13.0	647	7 BB655280	BB655280 BB655280
31	493.5	12.9	9534	14 DQ052790	DQ052790 Pan trogl
32	490	12.8	9022	14 AY402248	AY402248 Mus muscu
33	486.5	12.7	3600	6 CR859355	CR859355 Pongo pyg
34	486	12.7	2760	14 AY402758	AY402758 Mus muscu
35	486	12.7	3549	6 AK138211	AK138211 Mus muscu
36	485	12.7	2993	6 AK163818	AK163818 Mus muscu
37	477	12.5	1746	14 AY417765	AY417765 Homo sapi
38	476.5	12.5	631	7 BB612001	BB612001 BB612001
39	474	12.4	549	8 CN227517	CN227517 RCB020H02
40	468.5	12.2	4126	6 HSM807646	BM647500 Homo sapi
41	468	12.2	2217	6 AK041073	AK041073 Mus muscu
42	467.5	12.2	3180	6 AK028536	AK028536 Mus muscu
43	465	12.2	1752	14 AY411767	AY411767 Mus muscu
44	445	11.6	383	9 DR005740	DR005740 TC111897
45	445	11.6	2760	14 AY402757	AY402757 Pan trogl

ALIGNMENTS

RESULT 1
AK133486
LOCUS
DEFINITION
AK133486 3216 bp mRNA linear HTC 21-SEP-2005
Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
Riken full-length enriched library, clone:5031412M06
Product:VMA-like domains, matrilins and collagens, full insert
sequence.
ACCESSION
AK133486
VERSION
AK133486.1 GI:74225701
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
REFERENCE
1 Carinci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
2 Carinci, P. and Hayashizaki, Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159
JOURNAL
PUBMED
REFERENCE

AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komori, H., Akiyama, J., Nishii, K., Katsunari, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yonekura, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
	Riken Integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
PUBMED	11076861
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Komori, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishii, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Maranda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiya, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Guernicich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, U., Mazzarelli, J., Mombaur, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wymsh-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.
	Riken Genome Exploration Research Group Phase II Team and the FANTOM Consortium
CONSRM	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
PUBMED	11217851
AUTHORS	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Ose, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Balderelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapf, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brucic, V., Chotila, C., Corbani, L.E., Cousins, S., Dalle, E., Dargatz, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Guernicich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierecki, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagaishima, T., Numata, K., Okido, T., Pavan, M.J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravelet, J., Reed, J.C., Reed, J.D., Reid, J., Ring, B.Z., Ringwald, M., Saito, R., Schneider, C., Semp, C.A., Seton, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wymsh-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Komori, H., Nakamura, S., Sakazume, N., Sato, K., Shitaki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotoh, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shimokawa, K., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
	PANTOM Consortium
CONSRM	Analysis of the mouse transcriptome based on functional annotation
JOURNAL	Nature 420 (6915), 563-573 (2002)
PUBMED	12466851
AUTHORS	Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Fritth, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Bremner, S.E., Batalov, S., Forrest, A.R.,

AUTHORS	Zavolan, M., Davis, M.J., Wilming, L.G., Aldins, V., Allen, J.E., Banasi, I., Imamoto, A., Apweiler, R., Aizawa, R.N., Bailey, T.L., Banal, M., Baxter, L., Beisel, K.W., Berzano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffel, A., Clutterbuck, D.R., Crowe, M.L., Dally, E., Dalrymple, B.P., de Bono, R., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Heminger, P., Gingers, T.R., Gojobori, T., Green, R.E., Guernicich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ieko, K., Iwama, A., Ishikawa, T., Jakt, M., Kampin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollis, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, B., Madra, M., Marchionni, L., Maranda, H., Matsumura, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Motta, Tabar, S., Mulder, N., Nakano, N., Nakachi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlano, V., Pang, K.C., Pavan, W.J., Pavoni, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Saito, R., Schneider, C., Schonbach, C., Sekiguchi, K., Semp, C.A., Sero, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugita, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamashita, H., Zabarovsky, B., Zhu, E., T., Brucic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kato, S., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori, K., Sultana, R., Suzuki, M., Koch, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawai, H., Kawaguchi, N., Kawashima, T., Kojima, M., Kondo, S., Komori, H., Nakano, K., Niimura, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watanabe, A., Okamura-Ono, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.
	PANTOM Consortium
CONSRM	The transcriptional landscape of the mammalian genome
JOURNAL	Science 309 (5740), 1559-1563 (2005)
PUBMED	16141072
AUTHORS	Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakaniishi, M., Nakamura, M., Nishida, H., Yag, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Fritth, M., Ravelet, T., Pang, K.C., Hallinan, J., Matick, D., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Feghli, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.
	Riken Genome Exploration Research Group
CONSRM	Antisense transcription in the mammalian transcriptome
JOURNAL	Science 309 (5740), 1564-1566 (2005)
PUBMED	16141073
AUTHORS	8 (bases 1 to 3216)
	Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Komori, H., Murata, M., Nakamura, M., Niimura, N., Nishiyori, H., Nomura, K., Ono, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shitaki, T., Tagami, M., Tagami, Y., Waki, K., Watanabe, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL	Direct Submission
PUBMED	Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa, 230-0045, Japan (E-mail: genome-res@gs.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
AUTHORS	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
	Please visit our web site for further details.
REFERENCE	URL: http://genome.gsc.riken.jp/
REFERENCE	URL: http://fantom.gsc.riken.jp/

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: June 22, 2006, 12:29:33 ; Search time 349 Seconds

(without alignments)
5886.755 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826
Sequence: 1 LQEVAVSKETICKIKASAKM.....CKCRDQEGPHCENFLRRP 732

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xlh
-Oe/abs/ABSSMBE.spool/US10649591/runat_21062006_161715_8251/app_query.fasta_1
-DB=Issued_Patents_NA -OPMT=fastbp -SUFFIX=p2n.rml -MINMATCH=0.1 -LOOPTCL=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=humana0.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORMext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-HOST=abssmb04 -USER=US10649591 @CEN_1_1_204 @runat_21062006_161715_8251 -NCPu=6
-ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-NARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YCAPOP=10 -YCAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
- 7: /EMC_Celerra_SIDS3/ptodata/2/ina/PPCTUS_COMB.seq:*
- 8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
- 9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
- 10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2194.5	57.4	1837	3	US-09-312-283C-350 Sequence 350, App
2	509.5	13.3	3438	3	US-09-949-016-5648 Sequence 5648, Ap
3	509.5	13.3	3438	3	US-09-949-016-5649 Sequence 5649, Ap
4	505	13.2	3373	2	US-08-897-443-2 Sequence 2, Appl1
5	504	13.2	3449	3	US-09-907-794A-33 Sequence 33, Appl1
6	504	13.2	3449	3	US-09-905-125A-33 Sequence 33, Appl1
7	504	13.2	3449	3	US-09-902-775A-33 Sequence 33, Appl1
8	504	13.2	3449	3	US-09-906-700-33 Sequence 33, Appl1

9	504	13.2	3449	3	US-09-903-603A-33 Sequence 33, Appl1
10	504	13.2	3449	3	US-09-904-920A-33 Sequence 33, Appl1
11	504	13.2	3449	3	US-09-909-064-33 Sequence 33, Appl1
12	504	13.2	3449	3	US-09-905-381A-33 Sequence 33, Appl1
13	504	13.2	3449	3	US-09-906-618-33 Sequence 33, Appl1
14	504	13.2	3449	3	US-09-906-646-33 Sequence 33, Appl1
15	504	13.2	3449	3	US-09-904-462-33 Sequence 33, Appl1
16	504	13.2	3449	3	US-09-902-736A-33 Sequence 33, Appl1
17	504	13.2	3449	3	US-09-906-722A-33 Sequence 33, Appl1
18	504	13.2	3449	4	US-09-905-448-33 Sequence 33, Appl1
19	504	13.2	3449	4	US-09-903-562B-33 Sequence 33, Appl1
20	504	13.2	3449	4	US-09-906-679A-33 Sequence 33, Appl1
21	504	13.2	3449	5	US-09-907-841-33 Sequence 33, Appl1
22	499.5	13.1	3496	3	US-09-949-016-344 Sequence 344, App
23	492	12.9	3447	5	US-10-000-512-9 Sequence 9, Appl1
24	479	12.5	3026	5	US-10-000-512-7 Sequence 7, Appl1
25	477	12.5	2003	3	US-09-949-016-4469 Sequence 4469, Ap
26	477	12.5	2003	3	US-09-949-016-4470 Sequence 4470, Ap
27	451	11.8	3389	3	US-10-104-047-842 Sequence 842, App
28	426	11.1	8802	3	US-08-896-449A-1 Sequence 1, Appl1
29	426	11.1	8802	3	US-09-132-652-1 Sequence 1, Appl1
30	426	11.1	8802	3	US-09-886-900A-1 Sequence 1, Appl1
31	426	11.1	8802	3	US-09-662-478C-1 Sequence 1, Appl1
32	399	10.4	6153	2	US-08-347-594A-1 Sequence 1, Appl1
33	398	10.4	6153	3	US-08-463-682-2 Sequence 2, Appl1
34	396	10.4	8575	3	US-09-381-261A-2 Sequence 5433, Ap
35	386.5	10.1	1807	3	US-09-949-016-5435 Sequence 178, App
36	324.5	8.5	2773	3	US-09-991-181-178 Sequence 178, App
37	324.5	8.5	2773	3	US-09-990-444-178 Sequence 178, App
38	324.5	8.5	2773	3	US-09-997-333-178 Sequence 178, App
39	324.5	8.5	2773	4	US-09-992-598-178 Sequence 178, App
40	324.5	8.5	2773	4	US-09-989-735-178 Sequence 178, App
41	324.5	8.5	2773	5	US-09-989-726-178 Sequence 178, App
42	324.5	8.5	2773	5	US-09-997-528-178 Sequence 178, App
43	324.5	8.5	2773	5	US-09-989-718-178 Sequence 178, App
44	324.5	8.5	2773	5	US-09-997-343-178 Sequence 178, App
45	324.5	8.5	2773	5	US-09-997-653-178 Sequence 178, App

ALIGNMENTS

RESULT 1
US-09-312-283C-350
Sequence 350, Application US/09312283C
Patent No. 6573095
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murlson, James G.
APPLICANT: Kumble, Krishanand D.
TITLE OF INVENTION: Compositions Isolated from Skin Cells
FILE REFERENCE: 11000.1011c2
CURRENT APPLICATION NUMBER: US/09/312,283C
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 425
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 350
LENGTH: 1837
TYPE: DNA
ORGANISM: Mouse
US-09-312-283C-350

Alignment Scores:

Pred. No.: 8.93e-200 Length: 1837
Score: 2194.50 Matches: 434
Percent Similarity: 87.1% Conservative: 46
Best Local Similarity: 78.8% Mismatches: 68
Query Match: 57.4% Indels: 7
DB: 3 Gaps: 1

US-10-649-591-3 (1-732) x US-09-312-283C-350 (1-1837)

QY 1 LeuGlnGluValaHisValSerLySgluThrIleGlyLysIleSerAlaIleSerLyMet 20
 Db 196 CTTGAGGAAGTGCATGTGAACCGGAGACCATGGGGAAGATCGCTGTGGCCAGCAAAATTA 255
 QY 21 MetTrpCysSerAlaAlaValAspIleMetPheLeuLeuAspGlySerAsnSerValGly 40
 Db 256 ATGTGGGTCTCAGCCGGGTGCAATCTGTTTCTGTAGATGGCTCTCACACATCGGG 315
 QY 41 LysGlySerPheGluAspSerLyHisPheAlaIleThrValCysAspGlyLeuAspIle 60
 Db 316 AAGGAGACTTCGAGAGGTCCAGAGCGCTTGCCATCTGCTGTGATGCCCTTGAGCATC 375
 QY 61 SerProGluArgValArgValGlyAlaPheGlnPheSerSerThrProHisLeuGluPhe 80
 Db 376 AGCCCTGGCAGGGGTCAAGTCCGAGCGCTTGCAAGTTGGTTCACCTCTCATCTGGAAATTC 435
 QY 81 ProLeuAspSerPheSerThrGlnGlnGluValIleValIleAspGlyLeuAspMetValPhe 100
 Db 436 CCTTGGACTCTTCTCACTCAACAGAGTGAAGGAAGCATCAAGGGGATATGTTTTC 495
 QY 101 LysGlyGlyArgThrGluThrGluLeuAlaLeuIlePheTyLeuLeuHisAspGlyLeuPro 120
 Db 496 AAGGTGGGGCGCACCGAGACGGGCTTAGCCCTGAATA--CGCTTGAGCAGAGGGGTTTCCC 552
 QY 121 GlyIleArgHisAspValProGlnIleLeuIleIleValThrAspGlyLysSerGln 140
 Db 553 GAGGCGAGAAATGGCTGTGCGCCCAAGATCTTATGATGCTCAACGATGGCAAGTCCAG 612
 QY 141 GlyAspValAlaLeuProSerLyGlnLeuLysGluArgGlyValThrValPheAlaVal 160
 Db 613 GGGCCCGTGGCTCTCCCGGTAAAGCAGCTGAGAGAAAGGGGCGATCGTGGTGGCCGTA 672
 QY 161 GlyValArgPheProArgTrpGluGluIleLeuHisAlaLeuAspSerLysProArgGlyGln 180
 Db 673 GAGTCCGTTTCCAGAGTGGACAGAGCTGTCAAGTGGCCAGTAAAGCGGAAACCG 732
 QY 181 HisValLeuLeuAlaGluGlnValGluAspAlaThrAsnGlyLeuPheSerThrLeuSer 200
 Db 733 CATGTGCTGTGGCTGAGCAAGTGGAGATGCCAACCAATGGCTCTCTCAGACCCCTCAGC 792
 QY 201 SerSerAlaIleCysSerSerAlaThrProAspCysArgValGluAlaHisProCysGlu 220
 Db 793 AGCTCCGCACTGCAACCACTGCTGATCCAGACTGAGGGGTGAAACCTATCCCTGTGAG 852
 QY 221 HisArgThrLeuGluMetValArgGluPheAlaGlyAsnAlaProCysTrpArgGlySer 240
 Db 853 CGGAGGACGCTGGAGACCGTCAAGGAGCTGCTGGCAATGCCTTGTGTGGAGAGATCA 912
 QY 241 ArgArgThrLeuAlaValLeuAlaAlaHisCysProPheTySerTrpLysArgValPhe 260
 Db 913 AGGCAAGACAGACACTGTGGTGGCTGTGCCCTGTCTTACAGCTGGAAAGAGATGTC 972
 QY 261 LeuThrHisProAlaThrCysTyArgThrThrCysProGlyProCysAspSerGlnPro 280
 Db 973 CAGACACACCTGCGCAACTGCTACAGAACATGTGTCCAGGCCCTGTGATCTCCAGCCC 1032
 QY 281 CysGlnAsnGlyGlyTyThrCysValProGluGlyLeuAspGlyTyGlnCysLeuCysPro 300
 Db 1033 TGCCTAAATGAGGCGCGTGCATTCAGAAAGGTGTGAATGAGTACCACTGTCTTGGCCA 1092
 QY 301 LeuAlaPheGlyGlyGluAlaAsnCysAlaLeuLysLeuSerLeuGluCysArgValAsp 320
 Db 1093 CTGGCATTTGGAGGGGAAGTCAACTGTGCCCGAACTAGCTGGAATGCAAGATCGAT 1152
 QY 321 LeuLeuPheLeuLeuAspSerSerAlaGlyThrThrLeuAspGlyPheLeuArgAlaLys 340
 Db 1153 GTCCTCTTCTGTGAGCAGTTCTGAGGACACCACTTGGGGGCTTCCGAGGGGCAAG 1212
 QY 341 ValPheValLysArgPheValArgAlaValLeuSerGluAspSerArgAlaArgValGly 360
 Db 1213 GCCTTGTCAAGCGCTTGTGAGGCGGTGTGAGGAGGAACTCCGAGCCCGCGTTGGG 1272

QY 361 ValAlaThrTySerArgGluLeuValAlaVal-ProValGlyGluTyArgLysAsp-V 380
 Db 1273 ATAGCAAGATTATGGCAGGAATATTAATGTGGCGGTCCCTGTGCGGAGTACCAAGATTC 1332
 QY 380 AlaProAspLeuValTrpSerLeuAspGlyIleProPheArgGlyIleProThrLeuThrG 400
 Db 1333 TGCCTGACCTGATCAGAGACCTTTGACAGCATTCCTTTCAGCGGTGGCCGACCTTAACG 1392
 QY 400 LysSerAlaLeuArgGlnAlaGluArgGlyPheGlySerAlaThrArgThrGlyGln 420
 Db 1393 GAGGTGCTTGTCTCCAGGTGCGAGACCGGCTTGGAGTCCAGACGACTGTGTCAAG 1452
 QY 420 SPArgProArgArgValValIleLeuLeuThrGlnSerHisSerGluAspGluValAlaG 440
 Db 1453 ACAGGCGACGAGAGATGATGTTCTCTCATGMAATCAGCTCCAGATGAGGTCTTG 1512
 QY 440 LysProAlaArgHisAlaArgAlaArgGluLeuLeuLeuGluValGlySerGluAlaV 460
 Db 1513 GGCACAGACTCACGCAAGGGCTCGGAGACTACTCTCTGGGCGTGGCAGTGAATGCC 1572
 QY 460 AlaArgAlaGluLeuGluGluIleThrGlySerProLyHisValMetValTySerAsp 480
 Db 1573 TGCAGGCGGAGCTGTGAAGATCACCGGTAGCCGAAAGCATGTGATGTCCACACAGACC 1632
 QY 480 roGlnAspLeuPheAsnGlnIleProGluLeuGlnGlyLysLeuCysSerArgGlnArgP 500
 Db 1633 CTCAGGACCTGTTC-AGCCAAAT-CCAGAGCTCAGAGAGAGGCTATGCAAGCCAGCCGCG 1690
 QY 500 roGlyCysArgThrGlnAlaLeuAspLeuValPheMetLeuAspThrSerAlaSerValG 520
 Db 1691 CAGGCTGCCAGCAGACATGCACTGGACTGTCTTCTCTG--GATGCCCTGTGC-TGTGTGG 1748
 QY 520 LysProGluAsnPheAlaGluMetGlnSerPheValArgSerCysValAlaLeuGlnPheGluV 540
 Db 1749 GACGTGAAGCTTGTCCCAATGCAAGCTTATCAGAAATGCAACCTCGGTTTGAATG 1808
 QY 540 AlaAsnProAspValThrGlnValGlyLeu 549
 Db 1809 TGAATCTGATGTGACACAAATGGCTG 1837

RESULT 2

US-09-949-016-5648
 ; Sequence 5648, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CLO01307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5648
 ; LENGTH: 3438
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-5648
 Alignment Scores:
 Pred. No.: 4,47e-38 Length: 3438
 Score: 509.50 Matches: 200
 Percent Similarity: 38.3% Conservative: 126
 Best Local Similarity: 23.58 Mismatches: 316
 Query Match: 13.3% Indels: 210
 DB: 3 Gaps: 29

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2006, 13:14:36 ; Search time 2140 Seconds
(without alignments)
6304.590 Million cell updates/sec

Title: US-10-649-591-3

Perfect score: 3826
Sequence: 1 LOEVHVSKEITGKISASIKM.....CKCRDQWEGPHCENTRLRRP 732

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/abs/ABSSMB.spool/US10649591/runat_21062006_161725_8442/app_query.fasta_1
-DB=Published Applications NA Main -OPMT=fastlap -SUFFIX=p2n.rnpsm
-MNMATCH=0.1 -LOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blonum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcio -NORM=ext
-HEAPSIZE=500 -MNLN=0 -MAXLEN=2000000000 -HOST=ab8808
-USER=us10649591@cgn.1.1.1675 @runat_21062006_161725_8442 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPLICK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

1: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US09_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US09C_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10C_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10F_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10G_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10H_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10I_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10J_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US10K_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US11_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/pcdata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	3826	100.0	2810	7 US-10-295-027-123 Sequence 123, App

2	3826	100.0	2810	7 US-10-173-999-86 Sequence 86, App1
3	3826	100.0	2810	8 US-10-229-345-5 Sequence 5, App1
4	3826	100.0	2810	8 US-10-274-177-5 Sequence 5, App1
5	3826	100.0	2810	8 US-10-650-112-5 Sequence 5, App1
6	3826	100.0	2810	11 US-10-649-591-5 Sequence 5, App1
7	3826	100.0	2810	11 US-11-041-788-5 Sequence 969, App
8	3822	99.9	2268	7 US-10-295-027-969 Sequence 121, App
9	3799	99.3	2424	7 US-10-295-027-121 Sequence 970, App
10	3799	99.3	2424	7 US-10-087-080-22 Sequence 22, App1
11	3799	99.3	2424	7 US-10-173-999-84 Sequence 84, App1
12	3799	99.3	2424	7 US-09-930-020A-1 Sequence 1, App1
13	3799	99.3	3375	3 US-10-295-027-763 Sequence 763, App
14	3799	99.3	3375	7 US-10-295-027-776 Sequence 776, App
15	3799	99.3	3375	7 US-10-295-027-882 Sequence 882, App
16	3799	99.3	3375	9 US-10-702-180-1 Sequence 1, App1
17	3799	99.3	3375	9 US-10-450-763-17334 Sequence 17334, A
18	3887.5	88.5	3485	10 US-10-295-027-119 Sequence 119, App
19	3382.5	88.4	2481	7 US-10-173-999-82 Sequence 82, App1
20	3382.5	88.4	2481	7 US-09-866-050A-350 Sequence 350, App
21	2194.5	57.4	1837	6 US-10-152-661-350 Sequence 350, App
22	2194.5	57.4	1837	6 US-10-723-860-4635 Sequence 4635, App
23	927	24.2	352	9 US-10-027-632-97356 Sequence 97356, A
24	857	22.4	1992	6 US-10-027-632-97356 Sequence 3201, App
25	857	22.4	1992	6 US-10-027-632-97356 Sequence 7473, App
26	762	19.9	600	13 US-11-060-756-3202 Sequence 7473, App
27	762	19.9	600	13 US-11-060-756-7473 Sequence 7473, App
28	762	19.9	600	13 US-11-060-756-7474 Sequence 7474, App
29	762	19.9	600	13 US-09-918-715-217 Sequence 217, App
30	525	13.7	10558	7 US-10-096-534-14 Sequence 14, App1
31	525	13.7	10558	7 US-10-372-683-19 Sequence 19, App1
32	525	13.7	10558	8 US-10-734-864-24 Sequence 24, App1
33	525	13.7	10558	9 US-10-723-860-1064 Sequence 1064, App
34	525	13.7	10558	9 US-10-486-878-19 Sequence 19, App1
35	525	13.7	10558	10 US-10-956-157-1237 Sequence 1237, App
36	525	13.7	10558	10 US-10-287-436A-151 Sequence 151, App
37	525	13.7	10558	10 US-10-287-436A-764 Sequence 764, App
38	525	13.7	10558	10 US-09-814-553-19948 Sequence 19948, A
39	525	13.7	10558	6 US-10-084-817-43 Sequence 43, App1
40	525	13.6	10250	9 US-10-741-600-491 Sequence 491, App
41	521	13.6	10944	9 US-10-741-600-493 Sequence 493, App
42	521	13.6	10944	9 US-10-741-600-493 Sequence 493, App

ALIGNMENTS

RESULT 1
US-10-295-027-123
Sequence 123, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aiz, Nacasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevizi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394

```

/ PRIOR FILING DATE: 2001-11-15
/ PRIOR APPLICATION NUMBER: US 60/332,464
/ PRIOR FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: US 60/334,393
/ PRIOR FILING DATE: 2001-11-29
/ PRIOR APPLICATION NUMBER: US 60/340,376
/ PRIOR FILING DATE: 2001-12-14
/ PRIOR APPLICATION NUMBER: US 60/347,211
/ PRIOR FILING DATE: 2002-01-08
/ PRIOR APPLICATION NUMBER: US 60/347,349
/ PRIOR FILING DATE: 2002-01-10
/ PRIOR APPLICATION NUMBER: US 60/355,250
/ PRIOR FILING DATE: 2002-02-08
/ PRIOR APPLICATION NUMBER: US 60/356,714
/ PRIOR FILING DATE: 2002-02-13
/ Remaining Prior Application data removed - See File Wrapper or PAM.
/ NUMBER OF SEQ ID NOS: 1386
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 123
/ LENGTH: 2810
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-295-027-123

```

```

Alignment Scores:
Pred. No.: 0 Length: 2810
Score: 3826.00 Matches: 732
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

```

US-10-649-591-3 (1-732) x US-10-295-027-123 (1-2810)

```

QY 1 LeuGluGluValHisValSerIyGluThrIleGlyIleSerAlaIleSerIyMet 20
DB 158 CTCACGAGAGTCCATGTACCAAGAAACATCGGAAAGATTTCAGCTGCAGCAAAATG 217
QY 21 MetTrpCysSerAlaValAspIleMetPheLeuLeuAspGlySerAsnSerValGly 40
DB 218 ATGTGGTCTCGGCTGAGTGCACATCATGTTTCTGTAAATGGGTTCACAGCTCGGG 277
QY 41 LyeGlySerPheGluArgSerIyHisPheAlaIleThrValCysAspGlyLeuAspIle 60
DB 278 AAAGGGGAGCTTGAAGAGTCCAAAGCACTTGGCCATCAGCTGTGACGGTCTGACATC 337
QY 61 SerProGluArgValArgValGlyAlaPheGlnPheSerThrProHisLeuGluPhe 80
DB 338 AGCCCCAGAGGGTCAGAGTGGAGCAATTCAGTTCCACTCTCTCATCTGGAATTC 397
QY 81 ProLeuAspSerPheSerThrGlnGlnGluValIleValIleValIleValIleValIle 100
DB 398 CCTTGGATTCAATTTTCAACCAAGAGAGTGAAGCAAGATCAAGAGAGTGTTC 457
QY 101 LyeGlyValArgThrGluThrGluLeuLeuLeuIleValIleValIleValIleValIle 120
DB 458 AAAGGAGGGGCGACGAGACGGAACCTTGTCTGAAATACCTTGCACAGAGGTTGCT 517
QY 121 GlyGlyArgAsnAlaSerValProGlnIleLeuIleIleValIleThrAspGlyIleSerGln 140
DB 518 GGGAGCGAAGATCTTGTGTGCCCCAGATCTTCATCTGATCTGATGGGAAATGCCAG 577
QY 141 GlyAspValAlaLeuProSerIyGlnLeuIleGluArgGlyValThrValPheAlaVal 160
DB 578 GGGGATGTGCACTGCATCCAGCACTGAAGAAAGGGGTGCACTGTTCCTGCTG 637
QY 161 GlyValArgPheProArgTrpGluGluLeuHisAlaLeuIleSerGluProArgGlyGln 180
DB 638 GGGGTCAAGTTTCCAGGTGGAGAGCTCATGCACTGGCCAGCGAGCTAAGAGGGCAG 697
QY 181 HisValLeuLeuValGluGlnValGluAspAlaThrAsnGlyLeuPheSerThrLeuSer 200
DB 698 CAGGTCTGTGGCTGAGCAGGTGAGGATGCCACCAAGCGCTTTCAGACCTTCAGC 757

```

```

QY 201 SerSerAlaIleCysSerSerAlaThrProAspCysArgValGluAlaHisProCysGlu 220
DB 758 AGCTGCCCATCTGCTCTCAGCGCCAGCGAGTGCAGGCTCAGGCTCACCTGTGAG 817
QY 221 HisArgThrLeuGluMetValArgGluPheAlaGlyAsnAlaProCysTrpArgGlySer 240
DB 818 CACAGGACGCTGAGATGTGCTCCGGAGTTGCTGCGCAATCCCATGCTGGAGAGATCG 877
QY 241 ArgArgThrLeuValAlaLeuAlaHisCysProPheTrpSerTrpIlyAspValPhe 260
DB 878 CGGCGAACCTTCCGGTGGCTGCGACATGTCCTTCTACAGCTGGAAAGAGTGTTC 937
QY 261 LeuThrHisProAlaThrCysTrpArgThrThrCysProGlyProCysAspSerGlnPro 280
DB 938 CTAAACCACTCTGCCACCTGCTCAAGAGCACTGCCAGGCCCTGTGATCGCAGCCC 997
QY 281 CysGlnAsnGlyGlyThrCysValProGluGlyLeuAspGlyTrpGlnCysLeuCysPro 300
DB 998 TGCACAGATGAGGACATGTGTTCCAGAAAGACTGGACGGCTAACAGTCCCTGCCCCG 1057
QY 301 LeuAlaPheGlyGlyGluAlaAsnCysAlaLeuIleLeuSerLeuGluCysArgValAsp 320
DB 1058 CTGGCTTTGGAGGGAGGCTTAATCTGTGCTGAAAGCTGAGCTTGAAATCAGAGTGCAC 1117
QY 321 LeuLeuPheLeuLeuAspSerSerAlaGlyThrThrLeuAspGlyPheLeuArgAlaIlys 340
DB 1118 CTCCTCTCTCTGCTGAGACAGCTCTGGGGGACACACTGGACGGCTTCCTGCGGGCCAAA 1177
QY 341 ValPheValIlyAspPheValArgAlaValLeuSerGluAspSerArgAlaArgValGly 360
DB 1178 GTCTTGATGAAAGCGTTTGTGGCGCGCTGCTGAGGAGCACTCGCGGCCGAGTGGGT 1237
QY 361 ValAlaThrTrpSerArgGluLeuLeuValAlaValProValGlyGluTrpGlnAspVal 380
DB 1238 GTGGCACTATACAGAGAGAGCTGTGTGGCGGTGCTGTGGAGAGTACAGAGATGTG 1297
QY 381 ProAspLeuValTrpSerLeuAspGlyIleProPheArgGlyGlyProThrLeuThrGly 400
DB 1298 CTGACCTGTGCTGAGAGCTCGATGACATTCCTTCCGTGTGTGGCCCACTGACGGGCG 1357
QY 401 SerAlaLeuArgGlnAlaIleGluArgGlyPheGlySerAlaThrArgThrGlyGlnAsp 420
DB 1358 AGTGCTTGTGGGAGCGGAGAGCTGTGGCTTGGGAGCGCCACAGAGACGGCCAGAGAC 1417
QY 421 ArgProArgArgValValValLeuLeuThrGluSerHisSerGluAspGluValAlaGly 440
DB 1418 CGGCAAGTGAAGTGTGTGTGTCTCACTGAATCACTCCAGAGTGAAGTGTGCGGGC 1477
QY 441 ProAlaArgHisAlaArgAlaArgGluLeuLeuLeuGlyValGlySerGluAlaVal 460
DB 1478 CAGCGCGTACAGCAAGGGGGGAGAGCTGCTCTGTGGGTGTAAAGCAATGAGCGCTG 1537
QY 461 ArgAlaGluLeuGluGluIleThrGlySerProIyHisIleValIleValIleValIle 480
DB 1538 CGGCGAGAGCTGAGAGATCAACAGGAGCCCAACATGTGATGTCTACTCGATCTCT 1597
QY 481 GlnAspLeuPheAsnGlnIleProGluLeuGlnGlyIleLeuCysSerArgGlnAspPro 500
DB 1598 CAGGATCTGTTCAACCAAAATCCCTGAGCTGCAAGGAGAGCTGTGCAAGCCGAGCGGCA 1657
QY 501 GlyCysArgThrGlnAlaLeuAspLeuValPheMetLeuAspThrSerAlaSerValGly 520
DB 1658 GGGTGCAGGACCAACAGCCCTGAGCTCGTCTTCACTGTGGACACTCTGCTCAGTAGGG 1717
QY 521 ProGluAsnPheAlaGlnMetGlnSerPheValArgSerCysAlaLeuGlnPheGluVal 540
DB 1718 CCGAGAAATTTGTCTCAGATGCAAGACTTTGTGAAGAAGCTGTCCCTCCAGTTTGAGGTG 1777
QY 541 AsnProAspValThrGlnValGlyLeuValValIleValIleSerGlnValGlnThrAlaPhe 560
DB 1778 AACCTGAGTACACAGGTGGCTGTGTGTGTATGTGACGAGGAGTGCAGACTGCTTC 1837

```


D	3658	TACTGCAAGACGGCTACTGACACTGGTCTTCTGCTGGATGGCTCTCCAGGCTGTCCGAG	3717
Q	42	GIYSErPheGluAArgSerTyrNhrPheAlaLeuThrValCysAerGlyLeuAerLLeSer	61
D	3718	GCTAGATTGTAAGTGTGAAGGCTTTGTGTGGTGAACATGATGAGCGGCTGGCATCTCC	3777
Q	62	ProGluArgValAArgValGlyAlaPheGlnPheSerSerProHISLeuGluPhePro	81
D	3778	CAGAAAGTGGGCTCCGTGGCCGTGGTGGAGATACACACAGGCTTCCACGGCTTACATCCGG	3837
Q	82	LeuAerSerPheSerThrGlnGlnGlyValAlaArgLLeuYAsrMetValPheLys	101
D	3838	CTCAAGACCGGGAAGCCACGCTCAGACACTGGCGGCGCTGGCCAGCCAGGTGAATACCG	3897
Q	102	GIYGIYArg--ThrGluThrGlyLeuAlaLeuYAsrTyrLeuLeuNhrAArgGlyLeuPro	120
D	3898	GGCAGCCAGGTGGGCTCCACGACGAGGCTTTGAATATACACTGTTCCAA-----	3948
Q	121	GlyAlaArgAerAlaSerValProGlnLLeuLLeuLLeuAlaThrAerGlyLysSerGln	140
D	3949	-----ATCATCAGCAAGATCCACGCGCTGAAGCC	3978
Q	141	GIYAerValAlaLeu-----ProSerLysGln-----	149
D	3979	TTCTGTCATGCGCTCTGCTCTGATGGCCAGCCAGGACCCCAAGCAATGCCGGAATTT	4038
Q	150	-----LeuLysGluAArgGlyValThrValPheAlaValGIYValArg	163
D	4039	GTCGGCTACGTCCAGGGGCTGAAGAAGAAGAGTCTGTGATCCGGTGGGATTT---	4095
Q	164	PheProAArgTrpGluGlyLeuNhrAlaLeuAlaSerGluProAArgGlyGlnNhrValLeu	183
D	4096	-----GGGCCCCAGTCCCAACTCAAGCAATCCG-----CTC	4128
Q	184	LeuAlaGluGlnValGIYAerAlaLeuAerNhrAerNhrLysPheSerThrLeuSerSerSerAla	203
D	4129	ATTGAGAAGCAGGCGCCCTGAG-----AACAGGCC	4158
Q	204	LLeCysSerSerAlaLthrProAerPyrAerValGIYAerNhrValNhrProCysGlnNhrAArgThr	223
D	4159	TTCTGTCGAGCGGTGGTGGATGAG-----CTGAGACGACAAAG	4197
Q	224	LeuGluMetCValAArgGluPheAlaGIYAerAlaProCysTrpAArgLysSerAArgAArgThr	243
D	4198	GAAGAGATCTGTACCTACCTCTGTGACCTTGGCCCT-----GAAGCCCTCTCTCTACT	4251
Q	244	LeuAlaValLeuAlaAlaNhrCysProPheNhrYerTrpLysAerValPheLeuThrHis	263
D	4252	CTGGCC-----	4257
Q	264	ProAlaThrCysTrpAArgThrThrCysPheGluProCysAerSerGln-----Pro	280
D	4258	CCCCACATGGACAAAGTCACTGTGGGCGCGGGGCTTTGGGGCTTTGACGCTGGGGGCC	4317
Q	281	CysGlnAerNhrGlyLThrCys-----ValProGluGlyLeuAerGlyLysGln	296
D	4318	AAGAGCAACTCCATGGTTCTGATGTGGCGCTTTGTCCCGGAGCAATGGTGCACAAATT---	4374
Q	297	CysLeuCysPyrOleuAlaPheGlyGIYGIYAerNhrCysAlaLeuLysValLeuSerLeuGlu	316
D	4375	-----GGTAAACCCAG-----	4386
Q	317	CysAerValAerLeuLeuPheLeuLeuAerSerSerAlaGlyThrThrLeuAerGlyPhe	336
D	4387	-----TTT	4389
Q	337	LeuAlaGalaLysValPheValLysAerPheVal---ArgAlaValLeuSerGluAerSer	355
D	4390	AACAGAGCAAGGATTCATGAGAGGAGTATTCAGCGGATGAGTGTGGCCACAGACAGC	4449
Q	356	ArgAlaAArgValGIYValAlaLThrTrpSerAArgGluLeuValAlaValProValGIY	375
D	4450	---ATCCACGTCAAGGTGCTGCAAGTACTCTCAATGTCAGCCGTGGAGTACCTTTGACG	4506

QY 376 GlnUtyrGlnaerValProaerleuValTprserleuAerGlyIleProtheAerGly 355
 Db 4507 GAGGCACAGTCCAAAGGGGACATCTCTGCACGGGGTGGCGAGAGATCCGCTACAGGGCGGC 4566
 QY 396 ---ProthleuThGlySerAlaIleuArGlnAlaGluArGlyPheGlySerAla 414
 Db 4567 AACAGACCAACACTGGGGTGGCCCTGGGGTACCTCTGCACCAAGCTTC---TGGTGC 4623
 QY 415 ThrArGthrGlyGlnaerPProArGArGValValleuLeuThrGlnSerHisSer 434
 Db 4624 AGCCAGGGGTGACCGGGAGACGGCGCCCAACCTGGTCTACATGATCCAGAAATCTGGCC 4683
 QY 435 GlnAerGlnValAlaGlyProAlaArGHisAlaArGAlaArGlyIleuLeuLeuGly 454
 Db 4684 TCTATATGAGATC-----AAGAGCGTGGCTGGAGACATCCAGGTGGTGGCCATTGGA 4734
 QY 455 ValGlySerGlnAlaValArGAlaGluIleuGlnIleThrGlySerProIlyHisVal 474
 Db 4735 GTGGGCGCTTAATAGCCACAGTGCAGAGAGCTTGAGAGGATT---GGCTGGCCCAATGGCCCC 4791
 QY 475 MetValIyrSerAerProGlnaerPheanGlnIleProGluIleu---GlnGlyLys 493
 Db 4792 ATCTTCATCCACGAGACTTTTGAGACGCTCCCCGAGAGGCTTCGACCTGGTGTCTGCAGAG 4851
 QY 494 LeuCySerArGln-----ArGProGlyCyArGThr 504
 Db 4852 TGCTGCTCGGAGAGAGGGGCTGCAGATCCCACTCTCCCTGGCCCTGACTGC---AGC 4908
 QY 505 GlnAlaIleuAerPleuValPheMetleuAerPthrSerAlaSerValGlyProGluAerPhe 524
 Db 4909 CAGCCCTCGAGAGTATCTTCTCTGGATGGCTCTCCAGATTTCCACAGCTTCTTATTTT 4968
 QY 525 AlaGlnMetGlnSerPheValArGserCyValaIleuGlnPheGluValAerProAerVal 544
 Db 4969 GATGAATGAAGAGTTTGCTTAAGCTTTCATTCAAAAGCCAAATATAGGCTCATCTC 5028
 QY 545 ThrGlnValGlyLeuValValIyrGlySerGlnValGlnThralaPheGlyLeuAerPthr 564
 Db 5029 ACTCAGGGGTGAGTGTCTCCAGATGAAAGCATCACCAACATTGACGTCGATGAAAGCTGG 5088
 QY 565 LysPProthArGAlaAlaMetleuArGlnAlaIleSerGlnAlaProIlyrleuGlyVal 584
 Db 5089 GCCCGGAGAAAGCCCATTTGGTCTGAGCCTTGGAGACTATGCAGCGGGAGGGAGCGCC 5148
 QY 585 GlySerAlaGlyThrAlaIleuLeuHisIleIyrAerPlyAerValMetThrValGlnArGly 604
 Db 5149 AGCCAAATCGGGGAGTCCCTGGGGCTTGGCTGTCGCACTTGTGACTTCAGAAATGATGCT 5208
 QY 605 AlaArGProGlyValProIyAlaValValleuThrGlyGlyArGlyAlaGluAer 624
 Db 5209 GCGAGCGCGGAGCCTCAAAAGCGGTGTCATCTGGTCAAGCAGCCTCTGTGGANTCA 5268
 QY 625 AlaAlaValProAlaGlnIlyLeuArPheanAnGlyIleSerValIleuValIaGlyVal 644
 Db 5269 GTGATGACGACAGCTGATGCGCCGACAGTCCAAACAGATGACAGTGTTCCTATTTGGAATT 5328
 QY 645 GlyProValIleuSerGlnGly--LeuArGArGleuAlaGlyProArG----AerSer 661
 Db 5329 GGAAGATCGCTACAGTACAGCCAGAGTACCGAATCTTGGCAGAGCCCAAGCGACTCCAAC 5388
 QY 662 LeuIleHisValAlaAlaIyrAlaArPleuArGlyrHis-----674
 Db 5389 GTGGTAAGCTCAGCGAATCGAAGACTCTCC-TACCAATGTATCACTTGGGCAATTCCTT 5447
 QY 675 -----GlnAerValIleuIleGluTyrPleuCyS-----683
 Db 5448 CCTGCACAAAGTGTCTC-----TGAATTTGTATGAGATTCGATCGATGATGAGATGGAA 5501
 QY 684 ---GlyGluAlaIlyGlnProValIleuCyLysPProSerPro-----CysMe 699
 Db 5502 TGAGAAAGGCGCCGGAGAGCTGTGACCTTGGCAAGTGGCCACAGTGCACACCGTGACTTGGCA 5561